Journal
of Speech and Hearing
Research

March 1961 VOLUME 4 • NUMBER 1

Threshold of Hearing for Random Noise..... 3
RONALD HINCHCLIFFE

Parameters of Vowel Quality..... 10
GORDON E. PETERSON

Relationship between Auditory Comprehension and Word Frequency in Aphasia..... 30
HILDRED SCHUELL, JAMES JENKINS, AND LYDIA LANDIS

A Continuously Variable 360° Phase Shifter..... 37
COURTNEY STROMSTA AND WILLIAM L. DAWSON

Electrodermal Response Audiometry with Mentally Defective Children..... 41
JAMES W. MOSS, MARGARET MOSS, AND JACK TIZARD

An Articulation Test for Assessing Competency of Velopharyngeal Closure..... 48
HUGHLETT L. MORRIS, D. C. SPRIESTERSBACH, AND FREDERIC L. DARLEY

Pitch Discrimination and Articulation..... 56
RONALD K. SOMMERS, WILLIAM J. MEYER, AND ANN K. FENTON

Distance and Fingerspelling..... 61
HENRY M. MOSER, JOHN J. O'NEILL, HERBERT J. OYER, EDWARD A. ABERNATHY, AND BEN M. SHOWE, JR.

Low Sensation Level Delayed Clicks and Keytapping..... 73
RICHARD ALLEN CHASE, SAMUEL SUTTON, EDMUND P. FOWLER, JR., THOMAS H. FAY, JR., AND HOWARD B. RUHM

Maternal and Environmental Factors Related to Success in Speech Improvement Training..... 79
PHYLLIS BURGESS ANDERSLAND

Adaptation Effect and Spontaneous Recovery in Stuttering Expectancy..... 91
MARYANN PEINS

Call for Program Suggestions and Papers..... 100

Research News Notes.....36, 72, 78

The American Speech and Hearing Association

OFFICERS

President
G. Paul Moore, Ph.D.
Northwestern University

Executive Vice-President
Jack Matthews, Ph.D.
University of Pittsburgh

Vice-President
Duane C. Spriestersbach, Ph.D.
University of Iowa

OFFICERS-ELECT

President-Elect
James F. Curtis, Ph.D.
University of Iowa

Vice-President-Elect
Elise S. Hahn, Ph.D.
University of California
Los Angeles

COUNCIL

The Officers and the following Councilors:

Stanley Ainsworth, Ph.D. (1961)
Oliver Bloodstein, Ph.D. (1960-62)
William G. Hardy, Ph.D. (1960-62)
Ruth B. Irwin, Ph.D. (1960-63)
James Jerger, Ph.D. (1959-61)
Hayes A. Newby, Ph.D. (1960-63)
Mildred Templin, Ph.D. (1961-63)
Joseph M. Wepman, Ph.D. (1961-63)
Dean E. Williams, Ph.D. (1959-61)

EXECUTIVE SECRETARY

Kenneth O. Johnson, Ph.D.

The Journal of Speech and Hearing Research

EDITOR

Dorothy Sherman, Ph.D.

ASSISTANT TO THE EDITOR

Dorothy W. Moeller

STATISTICAL CONSULTANT

Leonard S. Feldt, Ph.D.

ASSOCIATE EDITORS

Oliver Bloodstein, Ph.D.
Arthur S. House, Ph.D.
D. E. Morley, Ph.D.
Hildred Schuell, Ph.D.
Arnold M. Small, Ph.D.
William R. Tiffany, Ph.D.
John C. Webster, Ph.D.
Joseph M. Wepman, Ph.D.

ASSISTANT EDITORS

Kenneth L. Moll, Ph.D.
Martin A. Young, Ph.D.

SPECIAL EDITORS

Ernest H. Henrikson, Ph.D.
Book Reviews

Martin F. Palmer, Sc.D.
Records

Elizabeth Moodie Prather, Ph.D.
Interlingua

BUSINESS MANAGER

Kenneth O. Johnson, Ph.D.

ASHA PUBLICATIONS BOARD

Virgil A. Anderson (1959-61); John Black (1961-63); Ernest J. Burgi, Representative, House of Delegates; Raymond Carhart (1960-62); Frederic Darley, Editor, *Monographs* (1960-62); Mary Huber, Editor, *Journal of Speech and Hearing Disorders* (1959-62); Kenneth O. Johnson, Business Manager of Publications and Editor, *Asha;* Dorothy Sherman, Editor, *Journal of Speech and Hearing Research* (1959-62); and Wendell Johnson, Chairman (1959-62).

Threshold of Hearing for Random Noise

RONALD HINCHCLIFFE

Measurements of the threshold of hearing for wide-band noise have been utilised in clinical audiology since Doerfler and Stewart (*3*) reported the use of a saw-toothed type of noise in a battery of audiometric tests designed to detect psychogenic deafness. There have, however, been no reports on the validity and reliability of random noise threshold measurements. This study was designed to explore that subject and, at the same time, to obtain data on the distribution of this function.

Procedure

Equipment. The random noise was produced by a masking generator that was constructed in the laboratory for use in conjunction with a speech audiometer. The instrument consists of a conventional thyratron circuit with an amplifier and a 3 db step variable attenuator, and has already been described (*10*). The signal was fed to a Standard Telephones and Cables Type 4026 earphone. This earphone was calibrated on a 3 cc coupler.

The study was conducted in the Medical Research Council's mobile hearing research unit (*9*) where the ambient noise level in the observer's booth conformed to the recommendations (*11*) of the Medical Research Council's Electro-Acoustics Committee.

Observers. A random sample population, stratified for age and sex, was drawn from a representative rural area in Great Britain. The population of this area had already been defined by private census for an unrelated survey. This survey area was the Vale of Glamorgan in South Wales. Audiologically speaking, the sample population was considered to be representative of the country as a whole because the pure-tone thresholds of hearing as a function of age for this population were found to be the same as those of another random sample population in south-west Scotland (*6, 8*).

The total sample drawn from the Vale of Glamorgan numbered 400, with ages ranging from 18 to 74 years inclusive, and with equal numbers of males and females. Of these, 100 people were drawn in the age group 18 to 24 years inclusive so that the size and limits of this subsample would be comparable to that of the sample on which the British Standard of the threshold of hearing by earphone listening was based (*1*). The other subsamples covered age spans of 10 years and there were 60 people in each subsample, with an equal number of males and females. Individuals in the sample were

Ronald Hinchcliffe (M.D., University of Manchester, 1955) was a member of the Scientific Staff of the Medical Research Council, England, when this study was done. Now in the United States on a National Institute of Health grant, he is Research Assistant Professor, Department of Otolaryngology, College of Medicine, University of Iowa.

TABLE 1. Number of individuals clinically and audiometrically examined in each age group.

Sex	*Age Group (in years)*						
	18 to 24	*25 to 34*	*35 to 44*	*45 to 54*	*55 to 64*	*65 to 74*	*All Ages*
Male	43	21	24	24	18	19	149
Female	37	25	17	20	21	19	139
Both	80	46	41	44	39	38	288

interviewed by home visitors who arranged to bring these people to the mobile unit for the hearing tests.

Experimental Method. At the mobile unit, a questionnaire relating to audiologic hazards was completed for each member of the sample. This was followed by a clinical otolaryngological examination and audiometric examinations for which methodology has been detailed previously (*9*). The air-conduction threshold of hearing of both ears of each observer was measured first with the pure-tone audiometer and then with what can be termed the random-noise audiometer. Pure-tone thresholds were measured at frequencies of 125, 250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 cps. The random-noise threshold was determined by the same psychophysical procedure of the method of limits, with the attenuator adjusted in 3 db steps.

To assess the reliability of, and the influence of 'practice' on, the random-noise threshold measurement, an effort was made to repeat the noise threshold measurement on a number of unselected observers. This second test occurred one to five weeks later.

Results

It was possible to perform the complete clinical and audiometric examination on 72% of the selected sample (the age and sex distribution of these observers is shown in Table 1) and to repeat the random-noise threshold determination on 26 unselected individuals.

For the purposes of the subsequent analysis, diseased ears were excluded on the basis of clinical criteria previously described (7). The distribution of the clinically normal ears is shown in Table 2.

Validity. The validity of the random-noise determination as a measurement of the threshold of hearing was assessed by comparing the threshold of hearing for random noise with that for individual pure tones. Product-moment correlation coefficients were calculated from the data on a sub-

TABLE 2. Number of clinically normal ears in each age group.

Sex	*Age Group (in years)*						
	18 to 24	*25 to 34*	*35 to 44*	*45 to 54*	*55 to 64*	*65 to 74*	*All Ages*
Male	73	36	33	36	23	24	225
Female	64	36	23	32	28	25	208
Both	137	72	56	68	51	49	433

TABLE 3. Range of thresholds of hearing for pure tones in audiograms used to calculate correlation of threshold of hearing for random noise with that for pure tones, when threshold shifts are due to presbycusis only.

Threshold Shift (in db)	*Frequency (in cps)*								
	125	*250*	*500*	*1000*	*2000*	*3000*	*4000*	*6000*	*8000*
Minimum	0	0	0	0	0	0	5	10	10
Maximum	30	25	35	50	45	50	45	65	80

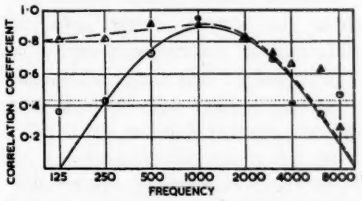

FIGURE 1. The correlation between the threshold of hearing for random noise and that for pure tones for observers with presbycusis only (dot within circle) and observers with conductive deafness (dot within triangle). The data for observers with presbycusis have been fitted with a sine curve which is defined by the equation: $r = \sin (88 \log f - 179) - 0.1$, where r = product-moment correlation coefficient, and f = frequency of pure tone. The horizontal dotted line marks the 95% significance level of the correlation coefficients.

sample (N = 20 ears) of women in the oldest age group. This age group was used because the effects of ageing would produce an appreciable scatter in the individual thresholds. The range of thresholds for each pure tone in this subsample is shown in Table 3. The calculated correlation coefficients, as a function of frequency, are shown in Figure 1. The data are fitted with a sine curve of the equation:

$$r = \sin (88 \log f - 179) - 0.1, \quad (1)$$

where r = the product-moment correlation coefficient for the threshold of hearing for random noise and for a pure tone of frequency f cps.

Fisher's z transformation showed that the predicted r was within the 95% confidence limits of the actual r.

For each frequency, a regression equation was derived to express the relationship of the random-noise threshold to the pure-tone threshold. Figure 2 shows the value of the gradient (a) and the intercept (b) in the equation:

$$y = ax_f + b \text{ db},$$

where y = threshold shift for random noise,

and x_f = threshold shift for a pure tone of frequency f cps.

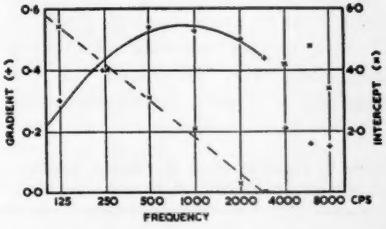

FIGURE 2. Values of the gradient (a) and the intercept (b) in the regression equation: $y = ax_f + b$, where y db = threshold shift for random noise, and x_f db = threshold shift for a pure tone of frequency f cps. Below 3000 cps the data for the gradient (a) have been fitted with a sine curve having the equation: $a = \sin (51 \log f - 60) - 0.45$, and that for the intercept (b) has been fitted with a straight line having the equation: $b = 13.8 - 4 \log f$ where f cps is the frequency of the pure tone.

TABLE 4. Range of thresholds of hearing for pure tones in audiograms used to calculate correlation of threshold of hearing for random noise with that for pure tones in cases of conductive deafness.

Threshold	*Frequency (in cps)*								
Shift (in db)	*125*	*250*	*500*	*1000*	*2000*	*3000*	*4000*	*6000*	*8000*
Minimum	10	5	10	10	0	0	5	10	20
Maximum	55	60	60	55	50	75	90	80	80

Figure 2 also shows that, for frequencies of 3000 cps and below, the data for the gradient (a) can be fitted by a sine curve with the equation:

$$a = \sin (51 \log f - 60) - 0.45, \quad (3)$$

and the data for the intercept (b) can be fitted by a straight line with the equation:

$$b = 13.8 - 4 \log f, \quad (4)$$

where f cps = frequency of the pure tone.

To assess the validity of measurements of the random-noise threshold in conductive deafness, ears giving a Rinne negative response to tuning fork tests were investigated. There were 11 ears in this category in the age group from 21 to 44 years. For this group of 11 ears, product-moment correlation coefficients were calculated for the threshold of hearing for random noise and that for pure tones. These coefficients are shown along with the coefficients for nonconductively deaf ears in Figure 1. The range of thresholds for each pure tone in this subsample is shown in Table 4.

Reliability and Practice Effect. For the 26 individuals who were tested twice, the test-retest reliability of the random-noise threshold gave a value of 0.89, in terms of a product-moment correlation coefficient. Student's t test was applied to these data to determine the practice effect on retest. A mean improvement of 1.9 db was significant at better than the 0.01 level.

Distribution. Table 5 shows, for each age group, the median thresholds of hearing for random noise, together with the upper and lower quartiles, and the

TABLE 5. Distribution of thresholds (median, upper and lower quartiles, and 95th percentile) of hearing for random noise as a function of age for clinically normal ears. Reference zero is median (Q_2) random-noise threshold for 18-to-24-year age group.

Age Group	*Thresholds (in db)*			
(in years)	Q_1	Q_2	Q_3	*95th %*
18 to 24	−1.6	0.0	2.4	5.9
25 to 34	−1.0	0.7	4.9	9.4
35 to 44	0.6	2.4	4.5	10.1
45 to 54	1.6	3.7	6.6	11.8
55 to 64	5.6	8.3	11.6	20.2
65 to 74	6.6	10.0	14.5	30.1

95th percentile. These data apply to clinically normal ears and the reference zero is the median random-noise threshold for the 18-to-24-year age group. Since there was no sex difference, the data for the two sexes have been combined.

Discussion

As indicated in Figure 1, for ears in which the effects of ageing only are apparent, the threshold of hearing for random noise shows the highest correlation with that for a pure tone of about 1000 cps, with decreasing correlation for frequencies above and below this tone. Pure-tone thresholds were analysed for the group on which the correlation coefficients were calculated. Threshold stimuli were expressed in terms of sound pressure levels for each of the nine frequencies used. The median threshold of hearing was lowest for a pure tone of 1000 cps, the 5th percentile was lowest for a pure tone of 2000 cps, and the 95th percentile was lowest for a pure tone of 500 cps. It was therefore expected that the correlation coefficient maximum would be found around 1000 cps, decreasing on either side of this. No particular importance is attached to the fact that a sine curve fitted the data.

With cases of conductive deafness (Figure 1), the correlations depart from the pattern of the more general case at the lower frequency end.

As indicated in Figure 2, when the threshold of hearing for random noise is computed from any pure-tone threshold of hearing, the noise threshold is always lower than the pure-tone threshold. This is because parameters of central tendency are being dealt with here. Examination of individual audiograms and the corresponding thresholds for random noise showed that the noise threshold was in each case the expected value when the absolute threshold of hearing was determined from the pure-tone thresholds by conversion of these to sound pressure levels.

The test-retest reliability ($r = 0.89$) of the random-noise threshold measurement is the same order as that which was found in the case of pure-tone audiometry, whether it be discrete frequency or continuous frequency audiometry (*2*). Determination of the threshold of hearing for random noise can therefore be considered to be highly reliable.

The mean improvement in the random-noise threshold, that is, about 2 db, occurring at the second test, is also of the same order as that previously demonstrated for pure-tone audiometry (*2*).

As evidenced by the threshold distribution information in Table *5*, there is a deterioration in the random-noise threshold with age. It has been shown previously that the deterioration in the thresholds of hearing for pure tones as a function of age follows an exponential course when the threshold is expressed in terms of a psychological magnitude scale (*5*), that is,

$$\log y = ax + b, \qquad (5)$$

where y = level of median threshold stimulus in terms of a psychological intensity scale,

x = chronological age in years,

and a and b are constants for a given frequency.

Ekman (4) has recently restated the psychophysical law so that it can now be written

$$\chi = k\ (\phi - \phi_o)^n, \qquad (6)$$

where χ = the subjective intensity,

ϕ = the physical intensity,

ϕ_o = the physical intensity at threshold,

and k and n are constants.

Equations (5) and (6) may be combined to obtain

$$\log\ (\phi - \phi_o) = ix + z, \qquad (7)$$

where ϕ = the physical intensity of the median threshold of hearing at x years of age,

ϕ_o = the physical intensity of the reference zero of the threshold of hearing,

and i and z are constants.

Calibration data for the random-noise audiometer indicated that the median threshold of hearing for people with normal hearing and in the age group 18 to 24 years corresponded to an over-all noise level of 27 db SPL (= 0.004477 dyn/cm²). If the position is accepted that the threshold shift due to ageing has commenced before the age of 21.5 years, then the value of ϕ_o for the purposes of this analysis is somewhat less than 0.004477 dyn/cm². Examination of the data showed that a suitable and probable value for ϕ_o is 0.004132 dyn/cm². Using this value, the median thresholds of hearing for random noise as a function of age have been replotted in Figure 3. The data conform very well to a straight line, confirming equation (7). The predicted median thresholds at a given age are within the 95% confidence limits of the actual data. Moreover, application of equation (7) to the data for pure tones also gives a good fit to the experimental data. Provided that adjustments are made in the value ϕ_o, a single straight line (that is, gradient and intercept constant and independent of frequency) can be shown to fit the experimental data for the frequencies 250 to 4000 cps inclusive. The differing rates of deterioration of pure-tone threshold with age may therefore be relative only and dependent upon expressing them on a particular scale. Moreover, the rate of deterioration of the threshold for these pure tones is the same as that for random noise. A difference in intercepts for the line applying to pure-tone threshold and that applying to the random-noise threshold is related to the difference between the width of the noise band and the width of the critical bands responsible for threshold stimulation.

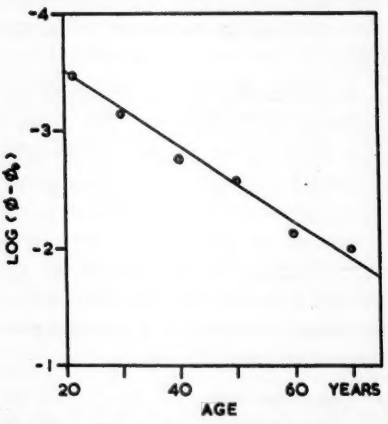

FIGURE 3. The random noise threshold as a function of age. The straight line fitted to the data has the equation: $\log\ (\phi - \phi_o) = 0.032\ x - 4.15$, where ϕ dyn/cm² = sound pressure of random noise that gives the median threshold stimulus at age x years, and $\phi_o = 0.004132$ dyn/cm².

Summary

The thresholds of hearing for random noise and for pure tones were measured on a random sample of a representative rural population. After the exclusion of clinically abnormal ears, analysis of the data showed that:

(a) there is a high correlation between the threshold of-hearing for random noise and that for pure tones of 1000 cps and of the two adjacent octaves;

(b) the test-retest reliability is high ($r = 0.89$);

(c) on the second measurement of the random-noise threshold, there is an improvement of about 2 db;

(d) as is also the case for pure tones, there is a deterioration of the random-noise threshold with age according to the expression: $\log (\phi - \phi_o) = ix + z$; where, in the case of the threshold for random noise, ϕ dyn/cm^2 = the physical intensity of the median threshold stimulus at x years of age; $\phi_o = 0.004132$ dyn/cm^2; $i = 0.032$, and $z = -4.15$.

Acknowledgment

The author is indebted to Dr. T. S. Littler, Director of the Medical Research Council Wernher Unit on Deafness, for his kind advice, to Prof. A. L. Cochrane and the Medical Research Council Epidemiological Unit for their assistance, and to Mr. G. C. Johnson for measuring the thresholds of hearing for random noise.

References

1. *British Standard for the Normal Threshold of Hearing for Pure Tones by Earphone Listening.* London: B.S. 2497, British Standards Institution, 1954.
2. Burns, W., and Hinchcliffe, R., Comparison of the auditory threshold as measured by individual pure tone and by Békésy audiometry. *J. acoust. Soc. Amer.*, 29, 1957, 1274-1277.
3. Doerfler, L., and Stewart, K., Malingering and psychogenic deafness. *J. Speech Dis.*, 11, 1946, 181-186.
4. Ekman, G., Two generalized ratio scaling methods. *J. Psychol.*, 45, 1958, 287-295.
5. Hinchcliffe, R., The pattern of the threshold of perception for hearing and other special senses as a function of age. *Gerontologia*, 2, 1958, 311-320.
6. Hinchcliffe, R., The threshold of hearing as a function of age. *Acustica*, 9, 1959, 304-308.
7. Hinchcliffe, R., The threshold of hearing of a random sample rural population. *Acta Otolaryng.*, 50, 1959, 411-422.
8. Hinchcliffe, R., Thresholds of hearing for pure tones as a function of age: An analysis in terms of psychological magnitude scales. Proc. Third Internat. Congr. on Acoustics, Stuttgart, in press.
9. Hinchcliffe, R., and Littler, T. S., Methodology of air-conduction audiometry for hearing surveys. *Ann. occup. Hyg.*, 1, 1958, 114-133.
10. Knight, J. J., and Littler, T. S., The technique of speech audiometry and a simple speech audiometer with masking generator for clinical use. *J. Laryng.*, 67, 1953, 248-265.
11. Medical Research Council, Committee on Electro-Acoustics, *Requirements and Technique of Usage of Pure Tone Audiometers.* London: Med. Res. Council, 1952.

Parameters of Vowel Quality

GORDON E. PETERSON

During the past few years there has been an increasing accumulation of evidence that formant frequencies are of primary significance in the identification of vowels. This evidence has come from theoretical studies of the transmission properties of the human vocal tract, from experiments involving systematic manipulations of the formant frequencies in studies of vowel quality, and from the successful development of instrumental procedures for the detection and generation of formant frequencies in speech analysis and speech synthesis, respectively.

The primary purpose of this paper is to present further data and considerations regarding the identifying parameters of spoken vowels. There is some evidence in the research literature that parameters other than the formant frequencies may influence human judgments of vowel qualities, and certain of these parameters are considered in the analysis of the data of the present study. The speech samples employed throughout were sustained vowels.

Gordon E. Peterson (Ph.D., Louisiana State University, 1939) is Professor of Communication Sciences, Director of the Communication Sciences Laboratory, and Chairman of the Program in Communication Sciences, University of Michigan. Much of the experimentation reported here was conducted while the author was a member of the technical staff of the Bell Telephone Laboratories. The manuscript was prepared while the author held a Summer Faculty Research Fellowship granted by the Horace H. Rackham School of Graduate Studies, University of Michigan.

Experimental Method

The data presented here are derived from spoken vowels which were analyzed by means of amplitude sections on the sound spectrograph. In this division of the paper a discussion of the techniques employed in collecting the data is followed by a review of the complications involved in formant measurements from such amplitude sections.

Vowel Matches. For the major portion of the study, speakers in the anechoic chamber at the Bell Telephone Laboratories listened to recorded reference vowels and attempted to produce vowels which were equivalent in phonetic quality. This is the technique employed in a research study previously reported (*17*).

The subjects were men, women, and children who were native speakers of English. They were trained to produce the non-English vowels, and the subjects did not seem to have greater difficulty in forming matches to the non-English than to the English vowels. Also, the children seemed as flexible in producing the vowels and as critical in making the judgments as the adults. In general, two different utterances were obtained for each speaker.

The instructions for obtaining the matched vowels varied from one subject to another, depending upon how readily the subject identified the concept of vowel quality, but in general subjects easily recognized the assigned task. Each subject and the experimenter worked together in the anechoic chamber to judge the recorded attempts by the subject. The work on any one vowel varied from a few minutes to as much as 30 or 40. In the experiment, both the speaker and the experimenter judged the speaker's recorded attempts to match the reference vowels, and almost without exception there was good agreement. Basically multidimensional judgments are involved. In some respects the procedure employed is similar to that required in psychophysical scaling (*23*). The vowels studied were oral (nonnasal) and only vowel matches which appeared to have normal phonation (normal vocal quality and natural fundamental voice frequency) were retained. It now appears, however, that the data might have been simpler to interpret if fundamental frequency also had been controlled.

Calibrations. The methods, instrumentation, and calibrations employed in the above mentioned study (*17*) also apply to those portions of the present report which involve the matching technique. Briefly, amplitude and frequency distortion measurements were made on the system with pure tones. The complete system employed for making the recordings and analyses (including the acoustical link) did not deviate significantly from a flat frequency response throughout the range of interest, so that formant amplitudes could be measured readily. Amplitude sections of a pulse tone were made approximately every hour throughout the experiment in order to calibrate the amplitude and frequency scales of the analyzer.

Formant Measurements. Difficulties in formant measurements when the fundamental voice frequency is high or when the formants are close together in frequency have been discussed previously (*19*). Lewis (*13*), van den Berg (*25*), and House and Stevens (7) have all shown that the formant bandwidths are generally less than 100 cps, whereas the fundamental voice frequency in speech normally exceeds 100 cps. Thus the voice harmonics do not provide an adequate sampling in the frequency domain for exact formant frequency, amplitude, and bandwidth determinations.

The chief instance discussed in the above indicated paper (*19*) is that in which $B_n > F_o$ (B_n = bandwidth of the nth formant, F_o = fundamental voice frequency, F_n = frequency of nth formant). Unfortunately this case is not particularly relevant to vowel formant measurements, for difficulties in such measurements result primarily from the fact that normally $B_n < F_o$; the measurement difficulties are augmented when $F_o > (F_n - F_{n-1})$.

Formant Data

In the experiment, the reference vowels to which matches were obtained lie in the areas of the vowel chart (*8*) normally symbolized by [i ɪ ɛ æ œ ɑ ɔ ʊ u ɯ ʌ ɝ]. The frequency and relative amplitude of each harmonic

of each of the matched vowels were tabulated, so that the formants could be specified. In some of these vowels, however, the formants lie close together in frequency and their measurement is often ambiguous. Thus the discussion in this paper has been restricted chiefly to vowels which have formants that are well separated in frequency. In this present section of the paper, data on the formant frequencies and relative amplitudes of such vowels are presented.

Formant Frequencies. It has been shown that the frequency positions of the vowel formants are not consistent in their absolute values from one speaker to another. It was found in a previous experiment that to a first approximation equivalent front vowels produced by different speakers tend to lie along lines of constant frequency ratio (*17*). An essential question is whether the formant ratios involve linear frequency relationships, or whether the nonlinear frequency characteristics of the ear affect listener judgment of formant ratios. The first two formants were chosen for examination since they are known to be very closely associated with vowel value, and since the second formant varies over a considerable range of frequency.

Figure 1 shows one conception of the use of auditory scales for an F_2 versus F_1 plot. Linear frequency scales were employed in constructing the graph, and, if extended, the straight dashed lines would pass through the origin, so that they are lines of constant frequency ratio. The location of these dashed lines and the range over which they extend are based on data previously published for the vowels indicated (*20*). The solid curves in this graph are lines of constant mel ratio (*24*).

Matches on six different vowels are shown in Figure 2. Generally, two independent amplitude sections were measured and plotted for each utterance by each speaker. Thus in most cases a speaker is represented by four points within the plot. Throughout Figures 2, 3, and 4, the reference vowels are shown by *x*. Dots and circles are alternated in order to provide a distinction among adjacent vowels. For any given vowel, points nearest the origin are for men and points farthest from the origin are for children. The dashed straight lines which pass through the origin represent

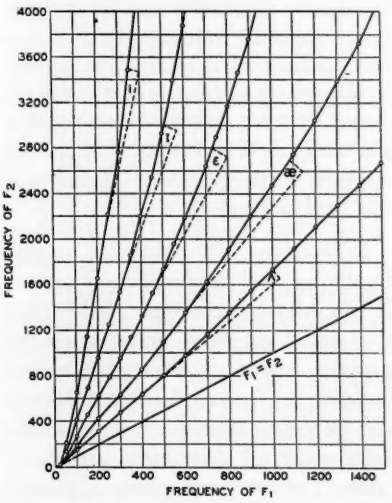

FIGURE 1. Ratio lines for the second formant versus the first formant for front vowels. The solid lines are for a constant mel ratio and the dashed lines are for a constant frequency ratio.

constant frequency ratios, and were drawn through an estimated distribution center for the various sets of points. The solid curved lines associated with the dashed lines are lines of constant mel ratio.

It will be noted that there is very little suggestion in Figure 2 that the distributions for [ʌ], [u], and [ɯ] lie along lines of constant frequency ratio. In this regard the distributions for [ʌ] and [u] appear to be consistent with the Peterson and Barney (*20*) data. It is evident from the graph that the data are entirely inadequate to determine in any substantial manner whether the formant ratios of matched vowels correspond more closely when expressed in frequencies or in mels. Thus at this time, the choice of scales for F_2 versus F_1 formant plots must be based primarily upon the resulting dis-

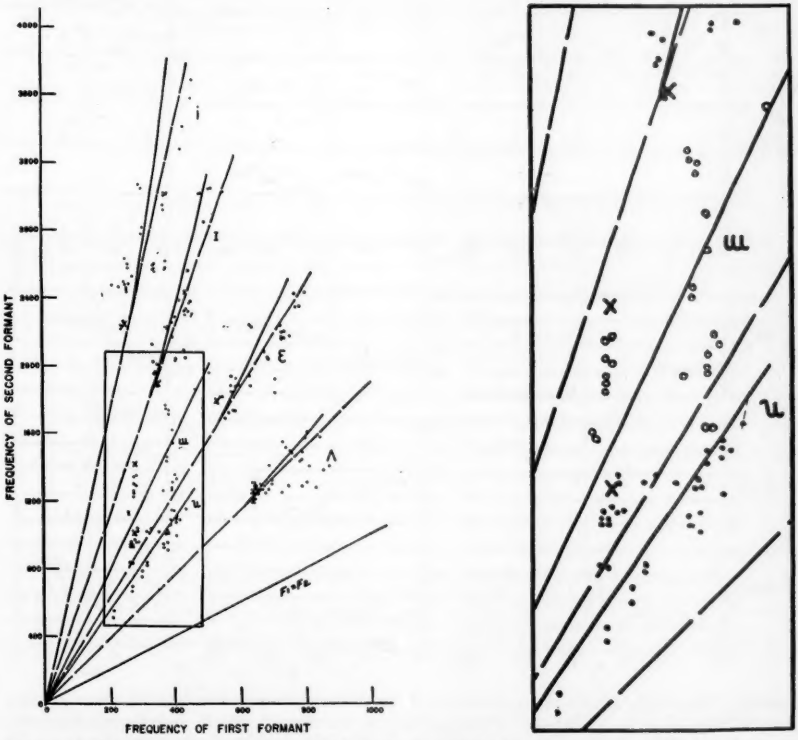

FIGURE 2. The frequency of the second formant versus the frequency of the first formant for a set of matched vowels by men, women, and children speakers. The references are shown by *x*. The dashed lines associated with the front vowels are for a constant frequency ratio and the solid lines are for a constant mel ratio. Dots and circles are alternated in order to provide a distinction among adjacent vowels. For any given vowel, points nearest the origin are for men and points farthest from the origin are for children. (Detail within rectangle enlarged at right.)

tribution (or spacing) achieved for the various vowels which it is desired to represent. Both the mel scale and the Koenig scale (*9*) seem to be fairly satisfactory. It may be noted, however, that when vowel loops for speakers with markedly different cavity sizes are plotted on log F_2 versus log F_1 paper, simple ratio lines for corresponding vowels are not preserved; also if the same linear scale is employed for both F_2 and F_1, the resulting vowel loop is considerably compressed in the dimension of the first formant. Linear scales which preserve formant ratio lines, and which the author has found convenient, are those employed for Figure 2, with the scale for F_1 expanded by a factor of two over the scale for F_2.

Analyses for three of the vowels,

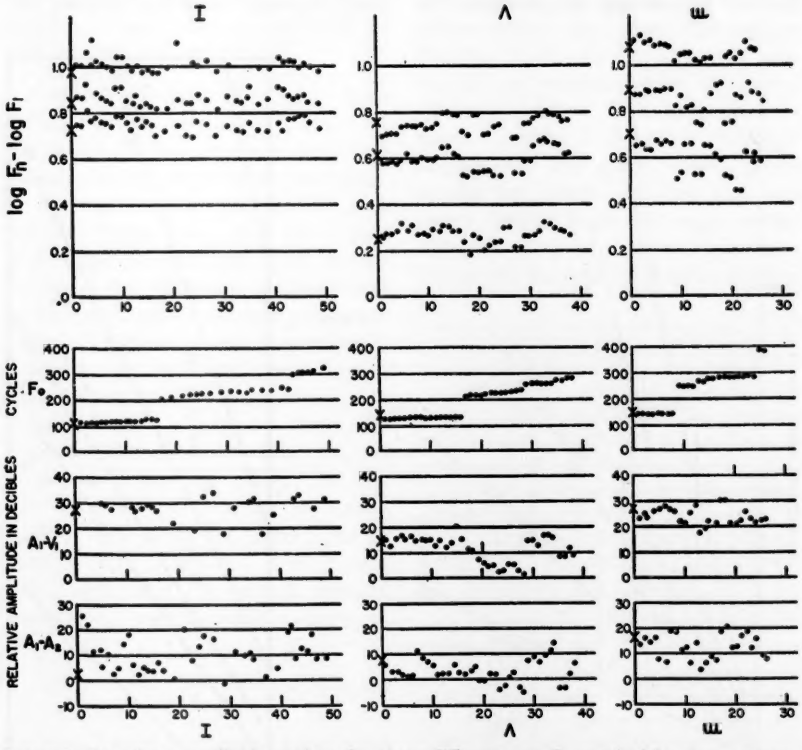

FIGURE 3. Data from amplitude sections for three different vowels matched by men, women, and children: log F_n — log F_1 (the logarithm of the second, third, and fourth formants, respectively, minus the logarithm of the first; F_o (fundamental voice frequency); A_1 — V_1 (the amplitude of the first formant minus the amplitude of the major valley or minimum between the first two formants in decibels); A_1 — A_2 (the amplitude of the first formant minus the amplitude of the second formant in decibels). The reference vowels are shown by *x*. Dots and circles are alternated in order to provide a distinction among adjacent vowels. For any given vowel, points nearest the origin are for men and points farthest from the origin are for children.

[ɪ], [ʌ], [ɯ], as shown in Figure 2, are presented in Figure 3. In the lower portion of Figure 3 are data for amplitude relationships, to be discussed subsequently. Corresponding data derived from the same spectrographic sections are shown near the center of the figure for fundamental voice frequency. Within each vowel the series is arranged according to the magnitude

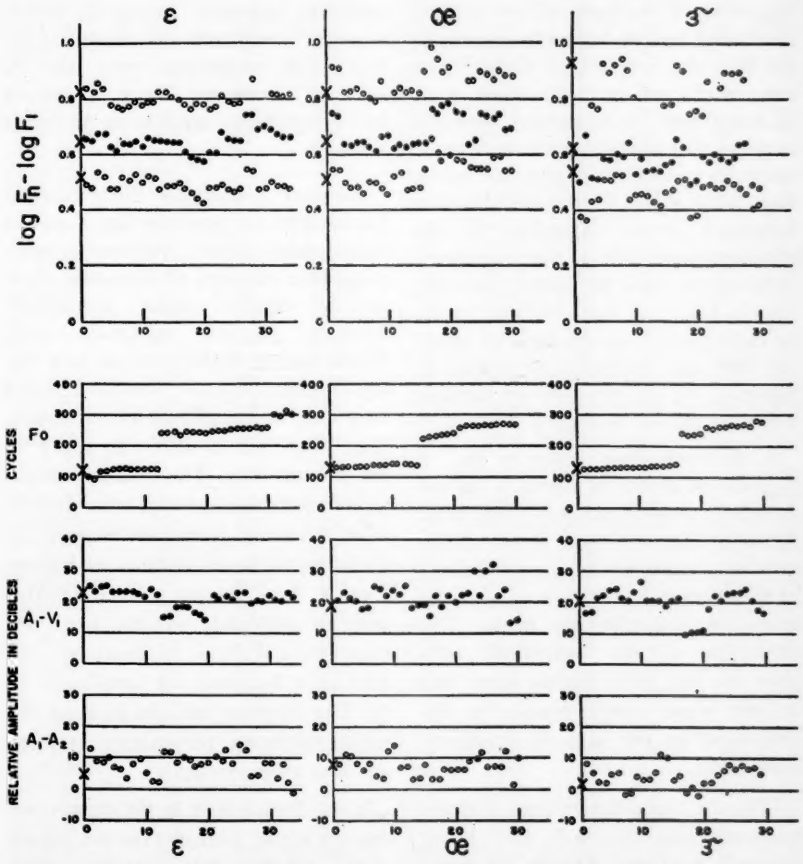

FIGURE 4. Data from amplitude sections for three different vowels matched by men, women, and children: log F_n — log F_1 (the logarithm of the second, third, and fourth formants, respectively, minus the logarithm of the first); F_o (fundamental voice frequency); A_1 — V_1 (the amplitude of the first formant minus the amplitude of the major minimum or valley between the first two formants in decibels); A_1 — A_2 (the amplitude of the first formant minus the amplitude of the second formant in decibels). The reference vowels are shown by *x*. Dots and circles are alternated in order to provide a distinction among adjacent vowels. For any given vowel, points nearest the origin are for men and points farthest from the origin are for children.

of the fundamental voice frequency. The upper portion of the figure was prepared to show the nature of the formant ratios which were obtained from the various utterances of the three vowels. This figure shows the logarithm of the ratio of the second, third, and fourth formants relative to the first; the first row of circles is for $\log_{10} F_2/F_1$ and so forth. Such a set of values may be considered direction tangents in a multidimensional formant space. In a logarithmic plot, the points align horizontally for identical formant frequency ratios, regardless of the absolute magnitudes of the frequencies involved. It will be noted that the vowels [ɪ], [ʌ], and [ɯ] can easily be distinguished on the basis of either the first two formants, as plotted in Figure 2, or on the basis of formant ratios, as shown in Figure 3.

A sequence chart similar to that of Figure 3 is shown in Figure 4 for the vowels [ɛ], [œ], and [ɔ]. This latter figure is for a set of three closely associated vowels which are difficult to discriminate on the basis of all of the acoustical measurements made. The similarities of the logarithmic ratio plots for the three vowels show that at least some vowels cannot be discriminated on the basis of frequency ratios alone.

When a three-dimensional formant frequency plot (F_3 vs F_2 vs F_1) was prepared for these vowels, however, it was found that the volumes encompassing each of the three sets of points were relatively isolated. That is, while the points fall near similar ratio lines, they differ in positions along those lines, that is, in absolute magnitudes. There is, of course, no means of reducing such a three-dimensional set of data to achieve better vowel separation than that obtained with the complete set. Various projections of such three-dimensional plots may be made, but optimum separation among all vowels is not achieved with any single projection. The projection onto the F_2 versus F_1 plane has, however, proved to be especially valuable in phonetics research.

Formant Amplitudes. Since formant bandwidths are generally less than the fundamental voice frequency, spectrographic displays of harmonic components cannot render particularly accurate formant amplitude data. While various techniques are now being developed for more precise formant frequency and amplitude measurements, these were not available to apply to the present data. The formant amplitude measurements which were derived from the vowel amplitude sections are shown in the lower portions of Figures 3 and 4. $A_1 - A_2$ represents the difference in amplitude of the first two formants, and $A_1 - V_1$ represents the difference between the amplitude of the first formant and the level of the major minimum (or valley) between the first two formants.

It was found early in the experiment that the higher formants for the vowels of the matching data were not as low in amplitude as those found in the Peterson and Barney (*20*) study. Careful and repeated checks of the calibrations established that the amplitude relations were indeed different from those of the previous data.

TABLE 1. Vowel fundamental, formant frequencies and amplitudes for a selected set of matched vowels, [i], [ɪ], [u], [ʊ]. Men, women, and children are indicated by M, F, and C, respectively.

Informant	F_0	F_1	F_2	F_3	F_4	A_1	A_2	A_3	A_4
For Vowel [i]									
REF(M)	125	250	2240	2740	3560	60	56	60	57
HKD(M)	122	260	2450	3180	3610	60	52	56	57
RHK(M)	123	260	2420	3200	3780	60	47	42	45
GEP(M)	125	250	2220	2880	3520	60	50	44	53
JB(F)	245	340	2700	3310	4310	60	50	53	58
TR(F)	246	340	2580	3100	4580	60	49	48	43
AC(F)	259	340	2560	3210	4010	60	56	50	51
JHP(C)	264	370	2900	3540		60	49	48	
GGP(C)	291	380	3000	3490	4250	60	50	49	48
JAG(C)	355	430	3370	4010	5320	60	45	44	43
For Vowel [ɪ]									
REF(M)	113	350	1890	2490	3380	60	58	53	56
RHK(M)	100	370	2050	2730	3750	60	35	42	37
AJP(M)	112	340	2020	2600	3540	60	55	55	57
GEP(M)	118	340	1940	2460	3420	60	54	54	55
HKD(M)	125	380	2220	2700	3740	60	55	55	53
JB(F)	212	440	2310	2940	4340	60	60	60	53
TR(F)	225	410	2030	2860	4220	60	46	50	42
GGP(C)	228	430	2420	3150	4610	60	43	47	43
JHP(C)	234	440	2550	3280	4610	60	51	47	48
JAG(C)	300	480	3000	3600	4800	60	49	46	35
For Vowel [u]									
REF(M)	140	280	980	2240	3500	60	46	26	35
GEP(M)	134	270	940	2140	3460	60	49	31	33
HKD(M)	137	290	1030	2190	3100	60	43	36	32
AC(F)	238	380	900	2380	3810	60	41	30	34
GEP(M)	239	310	810	2150	3820	60	40	28	38
HKD(M)	249	400	850	2190	3240	60	41	39	34
TR(F)	252	380	1010	2770	4030	60	52	35	37
JB(F)	269	400	1080	2610	4040	60	54	35	46
GGP(C)	277	390	1080	3080	4210	60	44	39	39
JHP(C)	290	410	1070	3190	4380	60	46	31	43
For Vowel [ʊ]									
REF(M)	131	470	1000	2450	3240	60	55	48	52
RHK(M)	127	480	1020	2100	3390	60	54	41	39
GEP(M)	128	470	960	2520	3230	60	51	38	42
HKD(M)	128	460	1080	2270	2850	60	48	45	37
EED(M)	136	490	1220	2560	3260	60	52	48	40
AC(F)	220	570	1060	2290	3520	60	58	42	41
TR(F)	234	470	1150	2570	3980	60	57	31	30
GGP(C)	263	500	1260	3180	4210	60	55	43	44
JHP(C)	264	500	1290	3480	4170	60	49	36	42
JAG(C)	282	560	1320	3780	4480	60	54	47	44

Tables 1, 2, and 3 present measurements of the fundamental voice frequency, the first four formant frequencies, and the first four formant amplitudes for a selected set of matches for each of the vowels studied. Table

TABLE 2. Vowel fundamental, formant frequencies and amplitudes for a selected set of matched vowels, [ɛ], [æ], [ʌ], [ɜ˞]. Men, women, and children are indicated by M, F, and C, respectively.

Informant	F_0	F_1	F_2	F_3	F_4	A_1	A_2	A_3	A_4
For Vowel [ɛ]									
REF(M)	127	530	1780	2410	3560	60	55	51	53
RHK(M)	96	560	1680	2500	3740	60	51	48	40
GEP(M)	125	580	1840	2500	3400	60	52	48	51
HKD(M)	125	580	1940	2560	3600	60	50	49	48
EED(M)	129	560	2300	2970	3740	60	58	53	51
RHK(M)	168	500	1830	2620	4030	60	55	51	47
TR(F)	233	630	1930	2800	4260	60	48	45	43
HKD(M)	235	610	1930	2440	3740	60	57	53	44
GEP(M)	237	570	1710	2370	3630	60	54	51	51
AC(F)	244	710	1900	2700	4370	60	52	49	47
JB(F)	246	740	2210	2950	4230	60	59	59	55
GGP(C)	250	720	2220	3400	4620	60	53	44	44
JHP(C)	253	630	2230	3490	4710	60	56	51	47
JAG(C)	295	770	2420	3780	5070	60	57	55	42
For Vowel [æ]									
REF(M)	124	710	1560	2470	3720	60	57	52	54
GEP(M)	125	700	1590	2400	3540	60	55	46	48
HKD(M)	128	760	1760	2560	3560	60	58	52	43
AJP(M)	128	760	1600	2430	3990	60	59	53	58
JB(F)	215	990	1850	2920	4300	60	58	53	50
JHP(C)	254	990	2140	3350	4320	60	53	48	44
GGP(C)	259	960	1970	2720	4020	60	59	50	42
TR(F)	260	910	2060	2880	3900	60	51	46	39
JAG(C)	280	1120	2380	3610	4760	60	57	46	37
For Vowel [ʌ]									
REF(M)	132	650	1270	2600	3630	60	53	48	49
HKD(M)	125	690	1260	2650	3440	60	55	49	51
RHK(M)	128	680	1240	2610	3660	60	54	44	46
GEP(M)	130	640	1260	2550	3600	60	58	43	45
EED(M)	132	660	1290	2720	4140	60	56	41	43
RHK(M)	163	440	1140	2580	3730	60	53	37	45
JB(F)	211	810	1400	2740	4200	60	56	45	47
AC(F)	215	640	1200	2260	4080	60	60	42	35
GEP(M)	218	650	1310	2640	3990	60	54	44	41
TR(F)	229	800	1300	2820	4100	60	57	43	38
HKD(M)	260	700	1220	2550	3640	60	56	53	39
GGP(C)	260	780	1430	3040	4420	60	52	43	40
JHP(C)	260	780	1480	3590	4680	60	54	45	48
JAG(C)	281	840	1570	3650	4950	60	54	45	37
For Vowel [ɜ˞]									
REF(M)	129	390	1340	1640	3300	60	58	51	39
RHK(M)	127	510	1360	1650	3050	60	58	53	32
GEP(M)	128	410	1320	1560	3460	60	55	53	39
HKD(M)	132	490	1410	1780	2980	60	57	53	39
TR(F)	226	380	1490	1760	4090	60	48	48	37
JB(F)	240	670	1610	2160	3840	60	58	46	51
GGP(C)	253	510	1570	1900	3800	60	55	51	33
JHP(C)	263	530	1680	2000	4210	60	53	50	40
JAG(C)	270	540	1620	2320	4180	60	53	48	35

TABLE 3. Vowel fundamental, formant frequencies and amplitudes for a selected set of matched vowels, [œ], [ɑ], [ɯ], [ɔ]. Men, women, and children are indicated by M, F, and C, respectively.

Informant	F_0	F_1	F_2	F_3	F_4	A_1	A_2	A_3	A_4
For Vowel [œ]									
REF(M)	136	480	1560	2110	3200	60	52	49	50
RHK(M)	131	470	1430	2080	3280	60	54	46	36
HKD(M)	133	490	1580	2060	3140	60	56	46	35
GEP(M)	137	490	1450	2060	3260	60	57	49	49
AC(F)	236	660	1650	2050	3870	60	55	55	45
GGP(C)	238	500	1830	2980	3930	60	54	46	40
TR(F)	248	420	1740	2180	4290	60	47	49	38
JHP(C)	262	520	1830	2280	3450	60	50	45	38
For Vowel [ɑ]									
REF(M)	125	750	1020	2620	3780	60	60	46	51
RHK(M)	100	730	1040	2300	3400	60	60	42	36
GEP(M)	125	710	1080	2530	3660	60	60	37	40
HKD(M)	125	850	1120	2750	3440	60	58	50	45
AC(F)	209	1090		2570	3970	60		39	30
JB(F)	208	1040		2810	4040	60		43	42
TR(F)	218	1130		3050	4140	60		42	35
GGP(C)	255	990		3060	3720	60		43	42
JHP(C)	259	1040		3630	4300	60		43	43
JAG(C)	271	1330		3770	4880	60		46	39
For Vowel [ɯ]									
REF(M)	141	280	1410	2210	3400	60	43	42	41
RHK(M)	138	260	1130	2040	3340	60	45	37	31
GEP(M)	141	280	1340	2220	3460	60	41	47	43
TR(F)	246	390	1260	2610	4110	60	49	39	37
JB(F)	250	380	1750	2450	4250	60	54	46	49
AC(F)	268	400	1340	2250	4100	60	57	37	43
JAG(C)	282	400	1610	3380		60	41	33	
GGP(C)	285	400	1140	2960	4280	60	43	34	40
JHP(C)	290	380	1450	2900	4380	60	49	36	47
For Vowel [ɔ]									
REF(M)	124	620	870	2630	3840	60	53	32	37
RHK(M)	91	540	840	2150	3230	60	56	31	34
HKD(M)	122	610	850	2680	3220	60	54	41	37
GEP(M)	125	610	890	2620	3720	60	53	32	36
EED(M)	125	660	880	2620	4050	60	55	35	43
AC(F)	222	800		2090	3550	60		43	41
JB(F)	235	820		2960	3760	60		37	45
TR(F)	230	740		2640	3910	60		31	31
GGP(C)	248	740		2830		60		30	
JHP(C)	261	780		3680	4620	60		45	47
JAG(C)	308	800		3660	4590	60		31	34

1 is for the vowels [i ɪ u ʊ], Table 2 is for the vowels [ɛ æ ʌ ɝ], and Table 3 is for the vowels [œ ɑ ɯ ɔ]. In constructing the tables the entire set of data was examined critically, and a characteristic match by each subject to each vowel was selected for presentation. For some subjects matches were

not available for all of the vowels, and for some of the vowels satisfactory measurements could not be made. For each vowel, the data are arranged according to the fundamental voice frequency of the particular match selected for presentation in the table. In general, vowel matches for which the frequency or amplitude formant data were obscure or ambiguous were omitted. In a few instances, gaps in the data indicate that a particular formant was not evident in the amplitude section. The first formant for each vowel was arbitrarily assigned a level of 60 db. The first formant almost universally had the greatest amplitude in the analyses. There were occasional amplitude reversals, but they rarely involved more than one or two decibels.

GEP was the reference (REF) speaker throughout. High pitch (sometimes falsetto) matches to three of the vowels [ɛ ʊ u] are included in the tables. These matches are not marked in any special way, except as they are indicated by repetition of informant initials and by the fundamental voice frequency. As discussed previously, the derivation of the frequencies and amplitudes of the first two formants for such vowels as [ɑ] and [ɔ] is difficult when the fundamental voice frequency is high. In such cases a single broad peak is usually evident in the amplitude sections. To indicate the general order of magnitude involved, a single frequency measure is given in the tables for the first two formants of these vowels. The first two formants were assigned the same peak amplitude (60 db).

One criterion in selecting subjects was that they should produce a relatively clear normal voice quality. It may be that both the reference speaker and the subjects employed somewhat sharper glottal tones and somewhat more open mouth positions than those common among a heterogeneous group of speakers. If so, these effects would explain at least some of the difference between the formant amplitude values in the tables and previously published data (*4*, pp. 109-120). It will be noted from both Figures 3 and 4 and also from Tables 1 and 2 that formant amplitudes have little to contribute to differentiation among the various vowels analyzed in this experiment.

Related Vowel Parameters

While formants provide the primary basis for distinguishing among vowels, the type of laryngeal excitation applied to the vocal cavities obviously affects formant detection and measurement in the output signal. Two extremes of phonation are whispering and falsetto. In the present experiment a number of matches were obtained to vowels produced with these two types of laryngeal excitation. Some informal observations on whispered vowels are presented below, but because of the complications in formant measurements, the results of the vowel matches are not presented in detail. The supralaryngeal modification of nasalization is another major aspect of vowel production which should be considered. Matches with nasalized vowels were not obtained, but some informal observations on nasalized vowels also are presented below.

Whispered Vowels. One basis for an interest in whispered vowels is that they have relatively continuous spectra. In such vowels the activation of the supralaryngeal cavities is produced with a partially closed glottis which creates a friction or random type of excitation. Because of the quasi-random nature of the output, more consistent information is derived if the energy composition of whispered vowels is evaluated over a period somewhat longer than that involved in normal short-time amplitude sections with the sound spectrograph. The long-time sectioner, previously described (*21*), provides a suitable integrating time, and it was applied to the analysis of a considerable number of matched whispered vowels. The analysis shows that there is a tendency for some resonance frequencies to shift upward when a vowel is whispered, and also that in many cases the low frequency energy (in the region of the first formant) is relatively weak in amplitude.

In Figure 5 wide-band sound spectrograms and narrow-band amplitude sections involving whispered speech are presented. In forming the vowels for these spectrograms a mechanical vibrator was held against the throat and an attempt was made to hold the vocal cavities in a fixed position throughout. A mirror was employed to check the position of the larynx. The left hand portion of each picture involves normal phonation with the vibrator turned off; at the discontinuity the vibrator was turned on. The combined weak level of low frequency energy from the glottis and the mechanical vibrator is evident in the amplitude sections. For each of the

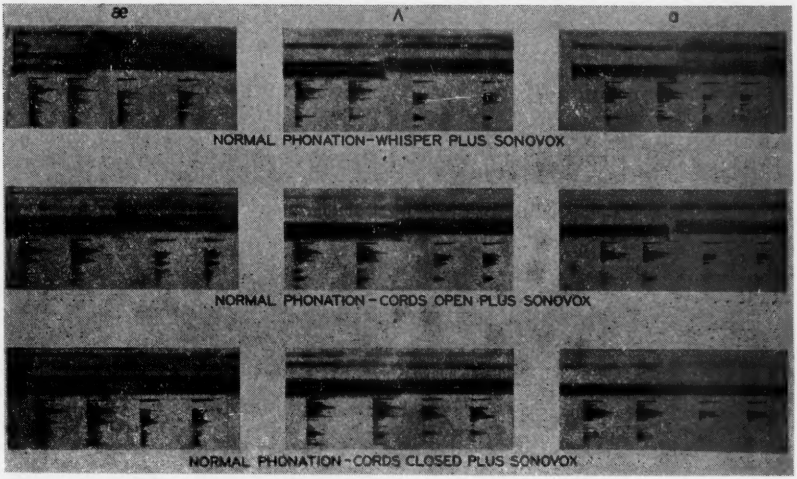

FIGURE 5. Broad-band sound spectrograms and amplitude sections of three different vowels, [æ], [ʌ], and [ɑ], produced under three different conditions of vocal cavity excitation. The first portion of each spectrogram is for normal phonation. The second part was produced as indicated, with the addition in each case of an external mechanical vibrator.

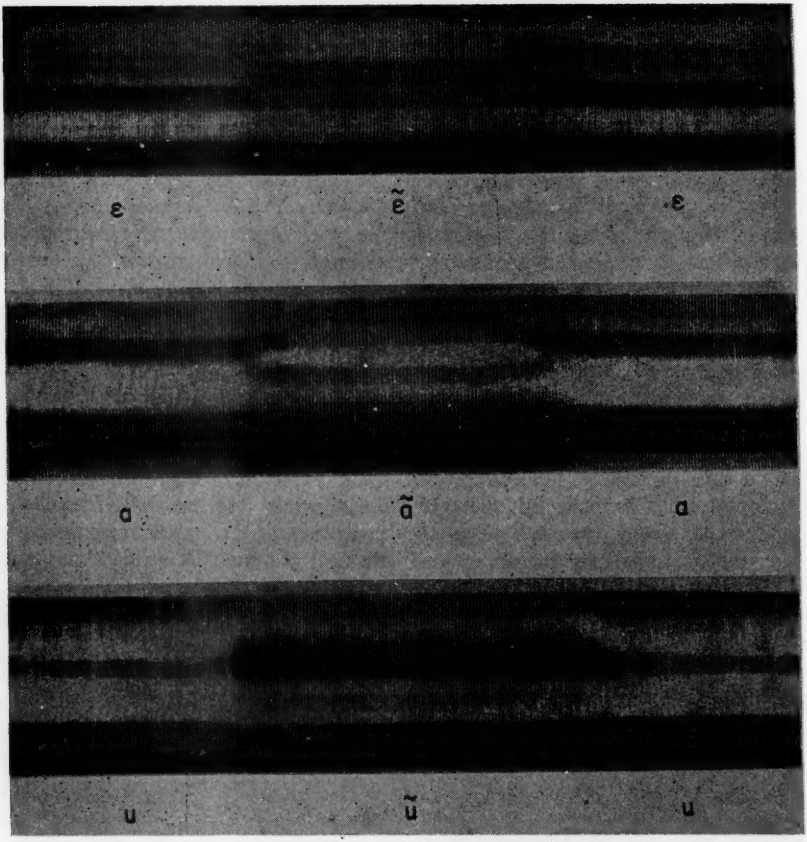

FIGURE 6. Broad-band spectrograms of an oral-nasal-oral sequence for the vowels [ɛ], [ɑ], and [u].

three vowels in the upper series, normal phonation was followed by whispering plus mechanical excitation. The shifts in resonance frequency positions are clearly evident. In the second series the normal phonation was followed by a relaxed or open glottis. Rather pronounced resonance shifts also appear in this series of spectrograms. By means of x-ray studies, Sovijärvi (22) found previously that there is a tendency to raise the tongue during whispered vowels, and associated changes in tongue position may account for some of the formant shifts evident in Figure 5. Following the normal phonation shown in the lower three spectrograms, the larynx was closed and the mechanical vibrator was turned on. In this latter case there is no observable shift of the resonance frequencies from the case of normal

phonation to that of closed glottis. This observation lends qualitative support to previous conclusions (*5*) that during normal phonation the larynx impedance is relatively high (as when the glottis is closed).

The above described effects in whispered speech are, of course, mechanical in nature. The difficulties in measuring the formant frequencies of whispered vowels made it impossible to determine from the vowel matching data whether minor adjustments are made in whispered vowels to render them perceptually equivalent to voiced vowels. It may be that the listener automatically adjusts to or corrects for those differences which result from applying a frictional glottal source to the same vocal tract configurations employed in phonated vowel formations. Possible secondary adjustments of formant frequencies during whispering to compensate for the absence of fundamental voice frequency have been discussed by Meyer-Eppler (*15*).

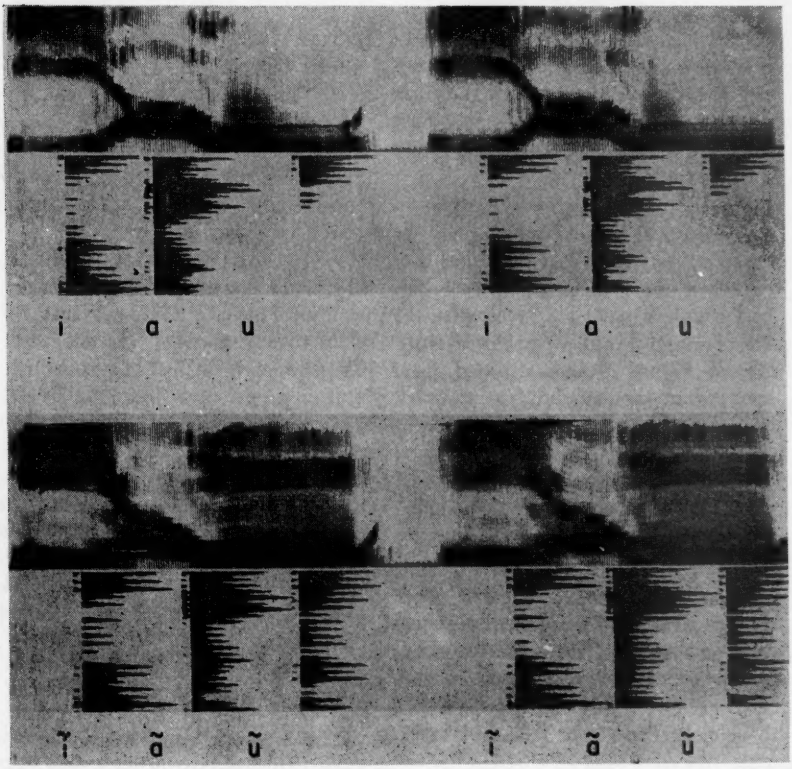

FIGURE 7. Broad-band spectrograms and amplitude sections for the vowel series: [i ɑ u]. The upper spectrogram shows two repetitions of the three vowels without nasalization, and the lower spectrogram shows two repetitions of the three vowels with the velum relaxed.

Nasalized Vowels. The discussion thus far has assumed a velopharyngeal closure. When there is an opening into the nasal passageway during vowel formations, the resulting effects upon perception may be considerable. Acoustically, the effects are primarily due to adding a branch cavity along the vocal tract; the acoustical theory of nasalization has been developed elsewhere (*3*, *6*) and will not be reviewed here. When there is a posterior opening into the nasal passageway, relatively marked effects upon the spectrum of the sound emitted from the mouth may be observed, whether the nose is open or closed anteriorly. An important consideration is whether the nasopharynx is coupled to the back cavity behind the hump of the tongue or to the front vocal cavity.

Figures 6 and 7 illustrate some of the acoustical effects which may be observed in nasalized speech. For Figure 6, three different vowels, [ɛ], [ɑ], and [u], were produced with the velopharyngeal sequence: closed-open-closed; Figure 7 shows broad band spectrograms and amplitude sections for two repetitions of the vowel series [i ɑ u] and [ĩ ɑ̃ ũ]. A mirror was employed to assist in holding other organs of the vocal tract, including the larynx, in as fixed a position as possible.

It can be seen from Figures 6 and 7 that generalizations about the acoustical spectrum of nasalized vowels must be made with considerable caution. The following effects appear to be associated rather frequently with nasal vowels.

a. There is a change in the level of the low frequency energy relative to the high frequency energy. The first formant may be broadened and shifted in frequency.

b. Secondary resonances may appear between the primary formants; these are most evident, of course, when the primary formants are well separated in frequency. The secondary resonances are more dependent upon the fixed nasal cavities, and thus are less variable than the primary formants. Numerical values for the secondary resonance frequencies cannot be given, of course, for like the primary formant frequencies they do not remain constant when different speakers produce similar vowels.

c. The energy level of the third formant may increase or decrease, and its frequency may shift up or down, depending upon the particular vowel formation. This fact is clearly illustrated in Figure 6.

It appears, then, that nasality involves a combination of several acoustical effects (*3*), and that these effects vary among different vowels. As in the case of many consonants, the unifying pattern appears to be associated with the vocal tract as the information source. Oral vowels appear to be somewhat simpler in their acoustical transformation than in their physiological form; but this is not true for the nasalization aspect of vowels, in which the degree of velopharyngeal closure is the essential component.

Vowel Perception

The above data indicate that neither a fixed formant frequency nor a formant frequency ratio hypothesis is adequate to explain vowel perception fully.

It appears that a relatively large number of factors must be considered in interpreting vowels. The following considerations appear to be of major significance.

Fundamental Voice Frequency. There are several indications that fundamental voice frequency may be a significant determinant of vowel quality. Since certain of the vowels displayed in Figure 2 do not appear to follow a formant frequency ratio pattern, it seems possible that fundamental voice frequency may have influenced the judgments when the vowel matches were made. It is possible, of course, either in speech synthesis or in human speech to alter the fundamental voice frequency without appreciably changing the formant frequencies.

In an experiment with a tone generator, Miller (*16*) showed that when only a portion of the vowel spectrum is presented, for example, only the first two formants, changes in fundamental voice frequency may considerably alter the judgment of vowel quality. In the back vowels, where a relatively pronounced effect of fundamental voice frequency upon vowel perception was observed, a number of factors may be involved. In certain of the back vowels the higher formants are relatively weak in amplitude, and experiments have shown that the intelligibility of back vowels is reduced very little when formants above the second are eliminated (*12, 18*). Since, in general, the formant frequencies are much higher for speakers with small cavity sizes, the first two formants are inadequate to distinguish back vowels among men, women, and children. For example, the first two formants of the [ɔ] of a child's voice correspond approximately to the position of the first two formants for the [ɑ] of a man. If a man raises his fundamental voice frequency to correspond to that of a child (falsetto), and the higher formants are removed by filtering, the acoustical result corresponds very closely to the [ɔ] of a child with low-pass filtering and may be so interpreted by a listener.

An even more general relation between fundamental voice frequency and vowel perception was observed by Wendahl (*26*). His experiments were conducted with a series resonance synthesizer, which simulated the first five formants and was excited by a recurrent impulse. Wendahl employed characteristic resonance positions for various vowels for both men and women speakers, and shifted the fundamental voice frequency over a wide range. His results showed that when the formant positions were held constant, the vowel value judgments varied with different fundamental voice frequencies.

Further research is much needed to determine the basic principles involved and to quantify the effects of fundamental voice frequency upon vowel perception. The possibility of an inverse effect in which the formant frequencies influence the perception of fundamental voice frequency should also be mentioned. Such an influence might be expected since the spectrum of a complex sound has an effect upon the perception of its fundamental frequency (*14*).

Formant Amplitude. Experiments with synthetic speech (*16*) have found an effect of formant amplitude upon

perceived vowel value. When formant amplitude relationships are sufficiently distorted, about the only two possibilities are that the sound will be perceived as some other vowel or that it will cease to sound like any human vowel. If a formant is reduced in amplitude so that it becomes lost in the background, or if it is increased in amplitude so that it obscures a nearby formant, then marked effects upon vowel perception may be expected. An interesting case is described by Delattre (*2*) in which a marked reduction in the amplitude of the first formant in synthesized speech causes vowels to be perceived as nasal.

In normal speech the major formants move about in the frequency domain with a continuity that would make them considerably more difficult to obscure by amplitude distortions than would be possible with synthesized vowels. The work of Fant (*4*, pp. 109-120) has shown that there are certain constraints on formant amplitudes. Altogether, it seems likely that in the perception of actual speech, formant amplitude variations are of secondary significance.

Phonetic Environment. Ladefoged and Broadbent (*1*, *11*) have demonstrated that the phonetic environment of an utterance may have a significant influence upon its perception. In their experiment a test word was preceded by a synthesized introductory sentence. In a series of presentations, the frequencies of the first or second (or both) formants were shifted in a systematic manner within the sentence which preceded the test word. It was found that the same recording of the test word, that is, the same physical stimulus, was judged differently depending upon the introductory sentence.

Such formant frequency shifts are primarily associated with speaker identification. While the experiments of Ladefoged and Broadbent involved synthesized speech, it is conceivable that similar effects may be observed in the interpretation of normal speech. In particular, the judgment of a vowel produced by a particular speaker within a context may perhaps vary, depending upon the nature and the extent of the utterance preceding the vowel. There are, in fact, several aspects of the nature and extent of environmental influence upon vowel production and perception which merit further investigation. It is obvious, however, that in general, linguistic judgments cannot be based upon environment alone, for if every sound were determined solely by its environment, there would be no means of determining any sound with which to construct an environment. The possible intereffects of phonetic environment upon vowel perception in the speech of individual speakers, however, pose an interesting challenge for future experimental linguistics research.

Conclusion

From the preceding data and discussion it appears that a number of factors are involved in vowel perception. Only to a first approximation do phonetically equivalent vowels have similar formant frequency ratios. The primary basis for this hypothesis is in the nature of acoustical models. Since the vocal cavities of men, women, and children are similar in general shape rather than in

size, it seems reasonable that vowels produced with similar articulatory shapes should have formant frequencies which do not correspond in absolute magnitudes, but which do have similar frequency ratios. Thus in the development of speech, for example, a child acquires certain articulatory neuromuscular coordinations. By employing a system of formant ratios, his articulatory patterns need not alter appreciably as his vocal cavities increase in size.

It should be emphasized that while phonetically equivalent vowels have somewhat similar formant frequency ratios, it does not follow that all vowels with similar formant frequency ratios are phonetically equivalent. For vowels to be perceptually equivalent it appears necessary that absolute values of their formant frequencies lie within certain limits. The influence of absolute magnitude of formant frequency is illustrated when vowel spectra are multiplied or divided by a constant factor, as by speed changes with a record-reproduce system (*10*). When the factor of multiplication or division becomes large, a marked reduction in vowel intelligibility results. Such an experiment has limited significance in the interpretation of vowel quality, however, since formant ratios remain constant, but formant bandwidths and the fundamental voice frequency are altered according to the factor of multiplication or division.

Formant amplitudes, fundamental voice frequency, and phonetic environment, in addition to formant frequencies, all appear to have an influence upon the perception of vowel qualities. This formulation seems to be consistent with more general considerations of perception. Characteristically, there are certain underlying patterns on which perceptual judgments are based. However, these patterns normally involve certain magnitudes and are associated with certain stimulus backgrounds or frames. When the magnitudes of the vowel parameters or when the frames in which the formant patterns are presented become sufficiently distorted, then different observer responses may result. According to such a pattern-experience concept of perception, markedly nonlinear effects may be anticipated when the stimulus parameter magnitudes fall far beyond their normal range.

Future Research

There yet remain many interesting and basic questions about the perception of vowels. Studies with human vowels are at present considerably handicapped by limitations in instrumental techniques for the measurement of vowel formants. Speech synthesis thus appears to provide a more satisfactory method for the investigation of certain of the problems.

Unfortunately, as yet there are very few basic psychophysical data on the perception of formant frequencies. The simplified analog of a vowel is a recurrently impulsed resonator; and although a great deal of experimental information is available about the auditory perception of pure tones, clicks, and random noise, little is known about the perception of recurrently impulsed resonators. Probably the most important problem for immediate investiga-

tion is the effect of fundamental excitation frequency upon resonator frequency discrimination. Other parameters of experimental interest are resonator bandwidth and stimulus amplitude and duration.

The relationships among fundamental voice frequency and the formant amplitudes and frequencies of perceptually equivalent vowel sounds is a subject yet requiring basic study. The perceptual evaluation of dynamic vowel patterns in speech is a third problem of major interest. The problem may be formulated as that of relating the moving formants of actual speech to discrete vowel patterns. The quantification of the intereffects of phonetic environment in vowel perception is a further major area for experimental study.

Summary

This study is concerned with the various acoustical parameters involved in vowel production and perception. Most of the data presented are based on analyses of sustained vowels which were matched to a set of reference vowels by various speakers in an anechoic chamber. Formant amplitudes, fundamental voice frequency, and phonetic environment are considered, in addition to formant frequencies, as factors which influence listener judgments of sustained vowels. A pattern-experience theory of vowel perception is suggested in which the ratios of formant frequencies may be considered to control judged vowel quality only within the normal limits of parameter magnitudes and the phonetic frame involved. Problems for future research on vowels are indicated.

References

1. Broadbent, D. E., and Ladefoged, P., Vowel judgments and adaptation level. *Proc. roy. Soc.*, 151, B., 1960, 384-399.
2. Delattre, P., Les attributs acoustiques de la nasalite vocalique et consonantique. *Studia ling.*, 8, 1954, 103-109.
3. Fant, C. G. M., *Acoustic Theory of Speech Production.* Stockholm: Roy. Inst. Technol., 1958.
4. Fant, C. G. M., On the predictability of formant levels and spectrum envelope from formant frequencies. In *For Roman Jakobson.* The Hague: Mouton ('S-Gravenhage), 1956.
5. Flanagan, J. L., Some properties of the glottal sound source. *J. Speech Hearing Res.*, 1, 1958, 99-116.
6. House, A. S., and Stevens, K. N., Analog studies of the nasalization of vowels. *J. Speech Hearing Dis.*, 21, 1956, 218-232.
7. House, A. S., and Stevens, K. N., Estimation of formant band widths from measurements of transient response of the vocal tract. *J. Speech Hearing Res.*, 1, 1958, 309-315.
8. International Phonetic Association, *The Principles of the International Phonetic Association.* Hertford: S. Austin, 1949 (available University College, London).
9. Koenig, W., A new frequency scale for acoustic measurements. *Bell Lab. Rec.*, 27, 1949, 299-301.
10. Kurtzrock, G. H., The effects of time and frequency distortion upon word intelligibility. Ph.D. dissertation, Univ. Illinois, 1956.
11. Ladefoged, P., and Broadbent, D. E., Information conveyed by vowels. *J. acoust. Soc. Amer.*, 29, 1957, 98-104.
12. Lehiste, I., and Peterson, G. E., The identification of filtered vowels. *Phonetica*, 4, 1959, 161-177.
13. Lewis, D., Vocal resonance. *J. acoust. Soc. Amer.*, 8, 1936, 91-99.
14. Lichte, W. H., and Gray, R. F., The influence of overtone structure on the pitch of complex tones. *J. exp. Psychol.*, 49, 1955, 431-436.
15. Meyer-Eppler, W., Realization of prosodic features in whispered speech. *J. acoust. Soc. Amer.*, 29, 1957, 104-106.
16. Miller, R. L., Auditory tests with synthetic vowels. *J. acoust. Soc. Amer.*, 25, 1953, 114-121.

17. PETERSON, G. E., The information-bearing elements of speech. *J. acoust. Soc. Amer.*, 24, 1952, 629-637.
18. PETERSON, G. E., Systematic research in experimental phonetics: 4. The evaluation of speech signals. *J. Speech Hearing Dis.*, 19, 1954, 158-168.
19. PETERSON, G. E., Vowel formant measurements. *J. Speech Hearing Res.*, 2, 1959, 173-183.
20. PETERSON, G. E., and BARNEY, H. L., Control methods used in a study of the vowels. *J. acoust. Soc. Amer.*, 24, 1952, 175-184.
21. PETERSON, G. E., and RAISBECK, G., The measurement of noise with the sound spectrograph. *J. acoust Soc. Amer.*, 25, 1953, 1157-1162.
22. SOVIJÄRVI, A., *Die gehaltenen, geflüsterten und gesungenen Vokale und Nasale der finnischen Sprache.* Helsinki: Druckerei-A.G. der finnischen Literaturgesellschaft, 1938.
23. STEVENS, S. S., The psychophysics of sensory function. *Amer. Scient.*, 48, 1960, 226-253.
24. STEVENS, S. S., and VOLKMAN, J., The relation of pitch to frequency: A revised scale. *Amer. J. Psychol.*, 53, 1940, 329-353.
25. VAN DEN BERG, Jw., Transmission of the vocal cavities. *J. acoust. Soc. Amer.*, 27, 1955, 161-168.
26. WENDAHL, R. W., Fundamental frequency and absolute vowel identification (abstract). *J. acoust. Soc. Amer.*, 31, 1959, 109-110.

Relationship between Auditory Comprehension and Word Frequency in Aphasia

HILDRED SCHUELL

JAMES JENKINS

LYDIA LANDIS

Previous studies have shown that frequency of usage in the language is an important factor in word recognition. Rosenzweig and Postman (*8*), reviewing the literature on perception and word frequency, reported that words presented individually in the presence of masking noise were identified more readily by normal subjects as frequency increased. Other studies have indicated that frequency is a significant factor in both auditory and visual recognition of words (*4*, *6*, *13*). Howes (*2*) found word frequency contributed to intelligibility of sentences. Siegel (*12*) found aphasic subjects reading orally, had less difficulty reading high frequency than low frequency words. Wepman, Bock, Jones, and Van Pelt (*16*), comparing the speech of an aphasic subject with that of a normal subject and a model of normal speech, concluded that all but the most frequent nouns, verbs, and adjectives tended to be absent from the speech of the aphasic subject.

Schuell and Jenkins (*11*) reported reduction of vocabulary in four language modalities in aphasia: comprehension of spoken words, naming, reading, and writing. The difficulty aphasic subjects have understanding spoken language has often been cited (*9*, *10*, *15*, pp. 144, 149, 463, 468). In order to explore the hypothesis that, in some ways at least, breakdown and recovery of language processes in aphasia are orderly and predictable in terms of what is known about language behavior in general, the present study explores the relationship between (a) ability of aphasic subjects to comprehend spoken words and (b) relative frequency of word usage in the language.

Hildred Schuell (Ph.D., University of Iowa, 1946) is Director, Aphasia Section, Neurology Service, Minneapolis Veterans Administration Hospital, and Clinical Associate Professor of Neurology, University of Minnesota. James Jenkins (Ph.D., University of Minnesota, 1950) is Professor of Psychology, University of Minnesota. Lydia Landis (B.A., Rockford College, 1957) is Member of Clinical Staff, Aphasia Section, Neurology Service, Minneapolis Veterans Administration Hospital.

Procedure

Subjects. Subjects were drawn from a total population of 154 aphasic patients in the Aphasia Section of the Neurology Service of the Minneapolis Veterans Administration Hospital. These patients had been tested on the Minnesota Test for Differential Di-

agnosis of Aphasia (*8*). These were consecutive admissions, except that patients were eliminated who were not aphasic, were unable to respond to testing, were psychotic, or showed regressive behavior. Within this series, subjects of the present study were 48 patients, 47 males and one female, to whom the Ammons Full-Range Picture Vocabulary Test, Form A (*1*), had been administered before treatment was initiated and again after treatment was terminated. Intervals between the two test administrations ranged from one to 12 months; the mean interval was 3.5 months. The age range of the 48 subjects was from 19 to 70 years; the mean age was 47. Education ranged from fourth grade through college, with a mean of 11 years of schooling. Of the total, 36 subjects were aphasic as a result of cerebrovascular accidents, six from trauma, three following infectious processes, and two from brain tumors. Etiology was undetermined for one patient.

Method. The Ammons Full-Range Picture Vocabulary Test, Form A (*1*), was administered in the standard manner. The word was spoken by the examiner, and the subject was asked to point to one of four pictures to indicate which picture best illustrated the meaning of the stimulus word.

The frequency of 83 of the 85 test words was plotted according to the General Count in the Thorndike-Lorge study (*14*, pp. 1-208). Two of the 85 words, *ebullience* and *dehydration*, were not used because they were not listed in the General Count. The 83 words were divided into four lists.

List I: 23 words which occurred 50 or more times per million words. The mean frequency of words on this list cannot be calculated from the Thorndike-Lorge list, since actual frequencies are not given but the best estimate is probably about 200 times per million (*14*, p. 249).

List II: 22 words which occurred from 49 to eight times per million words, with a mean frequency of 21.90.

List III: 22 words which occurred from seven to two times per million words, with a mean frequency of 3.77.

List IV: 16 words which occurred one or less than one time per million words, with a mean frequency of .81.

The number of errors, before and after treatment, on these 83 words was tabulated. Subjects were ranked in order of number of errors made on the initial test and divided into quartiles with 12 subjects in each, as follows: Q_1, 41 through 83 errors; Q_2, 29 through 38 errors; Q_3, 14 through 25 errors; and Q_4, 4 to 12 errors.

Results

The mean number of errors per subject was 29.58 on the initial test and

TABLE 1. Mean errors per word, before and after treatment, of 48 aphasic subjects on four word lists (I, II, III, and IV; 23, 22, 22, and 16 words, respectively) of decreasing frequency (occurrence per million: I, 50 or more; II, 49 to 8; III, 7 to 2; IV, one or less) from the Ammons Full-Range Picture Vocabulary Test; and results of *t* tests for evaluating differences between initial and final tests.

List	*Initial Test*	*Final Test*	*Difference*	t*
I	7.43	2.65	4.78	9.19
II	11.91	7.68	4.23	8.70
III	24.58	20.36	4.22	5.74
IV	27.88	23.06	4.82	6.95

*All obtained *t* ratios significant beyond the .001 level.

TABLE 2. Mean errors per word, before and after treatment, of 48 aphasic subjects on four word lists (I, II, III, and IV; 23, 22, 22, and 16 words, respectively) of decreasing frequency (occurrence per million: I, 50 or more; II, 49 to 8; III, 7 to 2; IV, one or less) from the Ammons Full-Range Picture Vocabulary Test. Means are presented according to quartile distribution of scores on initial testing (Q_1, 41 to 83 errors; Q_2, 29 to 38 errors; Q_3, 14 to 25 errors; Q_4, 4 to 12 errors).

Test	*Q*	*List*			
		I	*II*	*III*	*IV*
Initial	1	5.26	7.41	10.36	10.06
	2	1.39	3.18	7.68	8.75
	3	.61	1.14	4.45	6.13
	4	.17	.18	2.09	2.94
Final	1	1.48	4.00	8.36	8.31
	2	.87	2.73	6.95	8.19
	3	.30	.86	3.73	4.25
	4	.00	.09	1.32	2.31

21.81 on retest. The difference between the obtained means is significant beyond the .001 level ($t = 4.47$; $df = 47$). A positive change in scoring was made by 41 subjects, negative changes by five, and no change by two. The correlation between test and retest scales was .80.

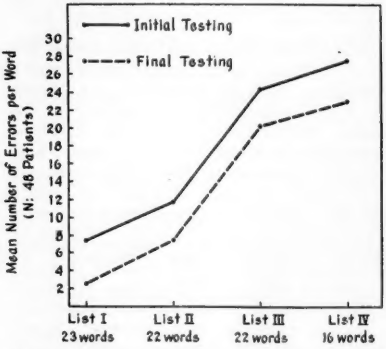

FIGURE 1. Mean number of errors per word made by 48 aphasic subjects on four word lists of decreasing frequency, on initial and final testing.

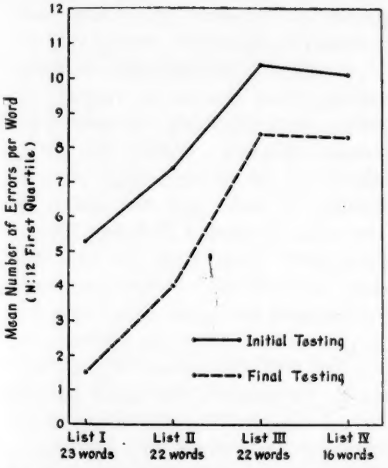

FIGURE 2. Mean number of errors per word made by 12 aphasic subjects in fourth quartile of initial distribution, on four word lists of decreasing frequency, on initial and final testing.

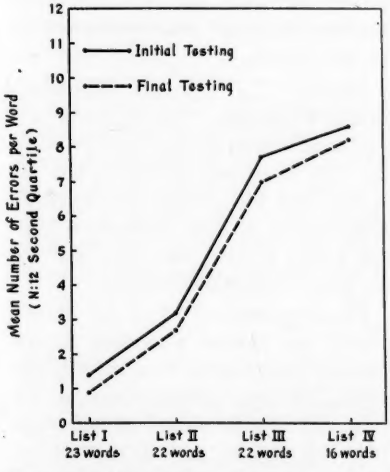

FIGURE 3. Mean number of errors per word made by 12 aphasic subjects in third quartile of initial distribution, on four word lists of decreasing frequency, on initial and final testing.

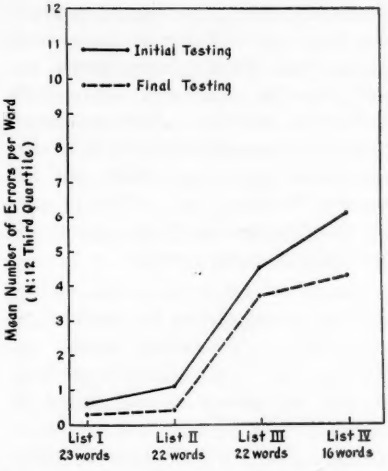

FIGURE 4. Mean number of errors per word made by 12 aphasic subjects in second quartile of initial distribution, on four word lists of decreasing frequency, on initial and final testing.

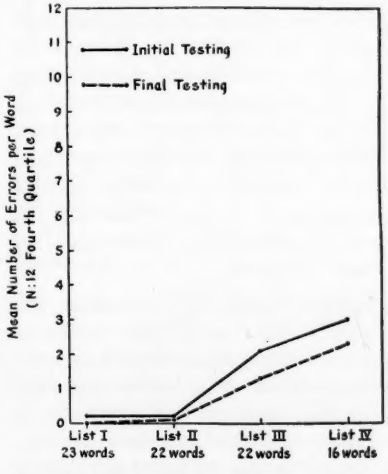

FIGURE 5. Mean number of errors per word made by 12 aphasic subjects in first quartile of initial distribution, on four word lists of decreasing frequency, on initial and final testing.

The mean number of errors per word on each of the four word lists is given in Table 1 and illustrated graphically in Figure 1.

Table 2 shows the mean errors per word for subjects in each quartile of the initial test distribution, for each of the four word lists on both test administrations. Figures 2, 3, 4, and 5 show the status of subjects in each quartile on each word list for both test administrations. These data support the hypothesis that word frequency is an important factor in the ability of aphasic subjects to comprehend spoken words. As frequency decreased, difficulties in comprehension increased. In other words, aphasic subjects had less difficulty comprehending words in common usage than words less commonly used. This appears to confirm, for the aphasic population used in this study, what Kirkpatrick and Cureton (*5*) found in normals on standard written vocabulary tests, namely, that frequency is an important determinant of difficulty of items when vocabulary is measured on a wide-range basis rather than in terms of precise word meaning. Findings in the present study also indicate definite gains in ability of aphasic patients to comprehend single words as recovery occurs.

Figures 2, 3, 4, and 5 show that regardless of severity of impairment of comprehension as measured by the Ammons test, low frequency words are the most difficult for all patients. Consideration of performance of subjects by quartiles showed that as total errors decreased, difficulty comprehending high frequency words decreased to approximately zero. As improvement occurred, more low

frequency words were comprehended. As would be predicted, subjects in each quartile made different kinds of gains in terms of frequency. Subjects in Q_4, who made the fewest errors on initial testing, made only negligible errors on Lists I and II in the first place, so that opportunity for gains on those lists was small, and most of the improvement necessarily came on less frequently used words. Conversely, patients in Q_1, who made the most errors on original testing, made the greatest gains on the most common words, and negligible gains on low frequency words. Individual subjects, on retest, tended to pass all the words they had passed before and, in addition, some words of adjacent lower frequency levels.

Means for subjects in the four quartiles maintained the same rank order on the four word lists on retest, indicating that subjects tended to remain in the same quartiles of the population after testing as before, although improvement was shown by most of them. The obtained r of .80 indicates that they also tended to maintain the same relative positions within quartiles.

Discussion

Word frequency appeared to be a significant factor in predicting word comprehension as measured by the Ammons test in aphasic subjects, in spite of conditions of the study which might be expected to obscure results. First, the Thorndike-Lorge General Count is based on frequency of written words, and the present study is based on comprehension of spoken words. Secondly, the Thorndike-Lorge Count was made almost 30 years ago; undoubtedly there have been changes in the frequency of usage of many words in this time. Thirdly, word length was not controlled in this study. Howes (*3*) found that in normal subjects auditory discrimination was better for long than for short words controlled for frequency. Fourthly, some of the pictures on the Ammons cards are clearer and less ambiguous than others.

Obtained errors were inconsistent with given frequencies for three words particularly. *Firecracker,* ranked according to Thorndike-Lorge frequency, was among the least frequently used words on List IV, yet only three subjects made errors on this word on the initial test, and none on the final administration. (It should be noted that Ammons classifies this word at grade level 2.7.) *Anaesthesia* (grade level 11.9, Ammons) was another word with fewer obtained errors than would be predicted from the Thorndike-Lorge Count. On the other hand, more subjects missed *numbers* (grade level 3.4, Ammons) than would be predicted from the frequency count. On the test card the correct response was a clock with numbers on the dial, but many patients pointed to a circle, perhaps considering it a zero, which may be an error particularly characteristic of aphasic responses.

Finally, there was a marked difference in the educational levels of patients within each quartile. Low educational levels undoubtedly depressed gains for many patients, particularly for those who made few errors on high frequency words on initial testing and had probably never been familiar with most of the low frequency words on the test.

It may be questioned whether or not previous exposure to the test cards and test words produced spurious retest gains. This, however, seems highly improbable with an aphasic population, even if the task had been simpler and the intervals between test administrations shorter.

The importance of the findings in relation to theory and treatment of aphasia would seem to be the following. First, there is a reduction of vocabulary, in comprehension as well as in speaking, of which the clinician should be aware. Secondly, this reduction is orderly; it is statistically closely related to the relative frequency of usage of words in the language, as is the performance of normal subjects on tests for perception and recognition of words. Thirdly, comprehension of words tends to improve during the course of treatment, also in an orderly and predictable manner. Lastly, the Ammons test appears to be a reasonably effective and sensitive indicator of such impairment and of improvement when it occurs, in aphasic subjects, and the information it provides can be utilized in selecting clinical materials of appropriate levels of difficulty for individual subjects.

Summary

The Ammons Full-Range Picture Vocabulary Test, Form A, was administered to 48 aphasic patients before and after treatment in the Aphasia Section of the Neurology Service of the Minneapolis Veterans Administration Hospital. Total errors on each of 83 words, and errors made by subjects on each quartile of the initial distribution, were analyzed for each test administration in relation to relative frequency of word usage shown by the Thorndike-Lorge General Count. Findings were as follows:

a. A significant gain was made between initial and final testing.

b. Subjects tended to maintain the same relative positions within quartiles. The correlation between initial and final tests was .80.

c. Relative frequency of word usage was an important factor in the ability of aphasic subjects to comprehend spoken words.

d. Gains made during recovery were related to both original performance level and word frequency.

References

1. Ammons, R. B., and Ammons, Helen S., *The Full-Range Picture Vocabulary Test.* Missoula, Mont.: Psychol. Test Spec., 1948.
2. Howes, D., The intelligibility of spoken messages. *Amer. J. Psychol.*, 65, 1952, 460-465.
3. Howes, D., On the relation between the intelligibility and frequency of occurrence of English words. *J. acoust. Soc. Amer.*, 29, 1957, 296-305.
4. King-Ellison, Patricia, and Jenkins, J. J., The durational threshold of visual recognition as a function of word-frequency. *Amer. J. Psychol.*, 67, 1954, 700-703.
5. Kirkpatrick, J. J., and Cureton, E. E., Vocabulary item difficulty and word frequency. *J. appl. Psychol.*, 33, 1949, 347-351.
6. Postman, L., and Adis-Castro, G., Psychophysical methods in the study of word recognition. *Science*, 125, 1957, 193-194.
7. Rosenzweig, M. R., and Postman, L., Frequency of usage and the perception of words. *Science*, 127, 1958, 263-266.
8. Schuell, Hildred, *The Minnesota Test for Differential Diagnosis of Aphasia.* (res.

ed.) Minneapolis: Univ. Minn. Print. Dept., 1955.

9. SCHUELL, HILDRED, Clinical observations on aphasia. *Neurology*, 4, 1954, 179-189.
10. SCHUELL, HILDRED, and JENKINS, J. J., The nature of language deficit in aphasia. *Psychol. Rev.*, 66, 1959, 45-67.
11. SCHUELL, HILDRED, and JENKINS, J. J., Reduction of vocabulary in aphasia. *Brain*, in press.
12. SIEGEL, G. M., Dysphasic speech responses to visual word stimuli. *J. Speech Hearing Res.*, 2, 1959, 152-160.
13. SOLOMON, R. L., and POSTMAN, L., Frequency of usage as a determinant of recognition thresholds for words. *J. exp. Psychol.*, 43, 1952, 195-201.
14. THORNDIKE, E. L., and LORGE, I., *The Teacher's Word Book of 30,000 Words*. New York: Teach. Coll., Columbia Univ., 1944.
15. WEISENBERG, T., and MCBRIDE, KATHARINE, *Aphasia, a Clinical and Psychological Study*. New York: Commonwealth Fund, 1935, and London: Milford, Oxford Univ., 1935.
16. WEPMAN, J. M., BOCK, R. D., JONES, L. V., VAN PELT, DORIS, Psycholinguistic study of aphasia: a revision of the concept of anomia. *J. Speech Hearing Dis.*, 21, 1956, 468-477.

► *RESEARCH NEWS NOTE*

Equipment for High Speed Cinefluorography

► Equipment for obtaining cinefluorographic pictures at camera speeds greater than the usual 24 frames per second has recently been obtained at the University of Iowa. This apparatus makes it possible to take films at camera speeds of 24 to 72 frames per second and to obtain a synchronized sound track. The sound track cannot be intelligibly reproduced by projection; however, it is used for visual identification of speech sounds.

The equipment consists of (a) a Pathé, variable-speed, 16 mm camera (type Webo M), (b) a Magnasync, magnetic film recorder (model 602C2), and (c) a Magnasync thyratron inverter circuit. The output from the inverter can be varied from 40 to 120 cps and drives both the recorder and the camera motors. An Electro-Voice, Model 644, unidirectional microphone is used to reduce the interference of equipment noise on the recording. An f:1.0 lens of one-inch focal length is used on the camera.

Before operation, synchronization marks are placed on the photographic film over the camera aperture and on the magnetic film over the record head. These marks are then used in transferring the magnetic recording to the photographic film by optical printing.

An experimental check on the picture-sound synchronization was made utilizing clap-sticks with lead inserts on the striking edges. At all film speeds the picture and sound track were found to be within one frame of exact synchronization. This degree of accuracy is considered to be more than adequate for research purposes.

Kenneth L. Moll, Ph.D.
Research Associate
University of Iowa
Iowa City, Iowa

A Continuously Variable 360° Phase Shifter

COURTNEY STROMSTA

WILLIAM L. DAWSON

Phase manipulation of audiosignal voltages, essential to the study of some problems which are the concern of the speech and hearing scientist, is generally in terms of phase shift so as to control the phase angle between two or more sinusoidal signals. To this end a method is chosen in conjunction with purpose to obtain a phase shift that is either (a) fixed or adjustable over a limited range, or (b) continuously variable over 360°. Since most textbooks devoted to electronics describe a number of circuits that may be adapted to obtain either a fixed or limited-range phase shift, the present discussion is limited to methods that permit a continuous shift of audiovoltage phases over a range of 360°.

A review of the literature disclosed three such methods. Everest (*4*) has described one method based on resistively mixing quadrature voltages. More specifically, the input signal is fed through two channels with one of the outputs being in phase to the input and the other shifted in phase 90° relative to the input by a resistance-capacitance (RC) network. The output of each channel is fed to a potentiometer having a continuous resistive element tapped every 90° and a brush-arm rotatable through 360°. The two potentiometers are ganged so that the two brusharms remain perpendicular to each other. As the shaft of the ganged potentiometers is rotated, two quadrature voltages are resistively mixed to give an output voltage which is dependent in phase on the position of the shaft. The inherent frequency dependency of the RC network requires an application of automatic volume control to supply a constant input voltage to the potentiometer associated with the network which in turn reduces amplitude fluctuations in the audiofrequency output.

A second method depends on the generation of a rotating magnetic field by an inductance goniometer. Such a field is produced by a Helmholtz coil arrangement (*1*, *2*) having a primary circuit composed of two vertical and two horizontal coils. One pair of coils is connected directly across the input whereas the remaining pair is connected across the input in series with an RC network chosen to assure that the current through the two pairs of

Courtney Stromsta (Ph.D., Ohio State University, 1956) is Assistant Professor, Departments of Speech, Anatomy, and Otolaryngology, Ohio State University. William L. Dawson (Assoc. E., Franklin University, 1950) is Technician, Department of Speech, Ohio State University. The work was supported by a grant from the National Institutes of Health, Division of Neurological Diseases and Blindness.

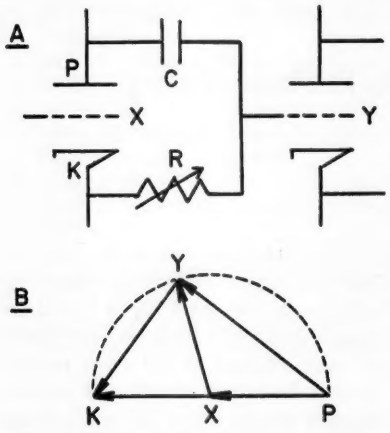

FIGURE 1. A, schematic diagram of one stage of the phase-shift circuit: P, plate; K, cathode; X, grid; Y, grid of following tube; C, capacitor; R, resistor. B, vector diagram of the phase-shift stage represented in A.

coils is essentially 90° out of phase. A secondary coil is pivoted in the center of the space between the paired-vertical and paired-horizontal coils so that it may be adjusted to any desired position relative to the primary coil. The rotating magnetic field formed by the currents in the primary coils induces a voltage in the secondary coil, and the phase angle between the output and the input voltages is dependent upon the position of the secondary coil. With this method RC values must be chosen as a function of frequency to minimize amplitude fluctuations of the output voltage as phase is shifted through 360°.

Mahren (*5*) has described a third method that introduces a phase shift in the fixed frequency of a beat-frequency oscillator by tuning an associated circuit through resonance. The phase shift appears in the variable audiosignal output after heterodyning. As stated by the author, two advantages of the method are (a) the phase shift is independent of frequency, and (b) no amplitude fluctuations occur as phase is shifted through 360°. A modification of the method has been described by Dolch (*3*) in which phase shift is accomplished by capacitive mixing of quadrature voltages.

The purpose of the ensuing discussion is to describe an additional method of shifting audiovoltage phases continuously through 360°. The unit in which the method is applied is both simple and economical to construct. It is thought that it circumvents in part the limitations or the expense involved in obtaining continuous phase shift by the three methods described above.

Theory of the Method

The operation of one stage of the phase-shifting unit as to phase and amplitude of the signal is described with regard to Figure 1. Vectors PX and XK of Figure 1B represent the two equal potential differences from plate (P) to grid (X) and from grid (X) to cathode (K) of Figure 1A. In like manner, vectors PY and YK of Figure 1B represent the potential differences from plate (P) to grid (Y) and from grid (Y) to cathode (K) of Figure 1A. It is then evident that vector XY represents in magnitude and phase the potential on grid (Y).

In Figure 1B the triangle PYK is a right triangle and X is the origin of a circle whose radius would intersect points P, Y, and K; therefore, vectors XY, PX, and XK are equal in magnitude. The potential represented by vector XY changes in direction (phase) but not in magnitude (amplitude).

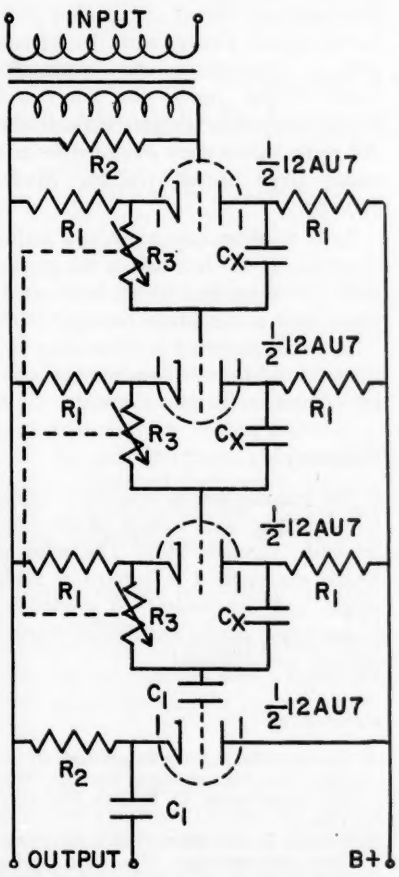

FIGURE 2. Schematic diagram of the continuously-variable 360° phase shifter. R_1, 5.6K ohms; R_2, 47K ohms; R_3, 100K ohms potentiometer (ganged); C_1, .05 μf; C_x, .03, .01, and .005 μf mounted on each section of a three-section, three-position rotary switch.

Stated otherwise, in each of the three phase-shifting stages of the unit shown schematically in Figure 2, the combination or vector summation of resistance (R_3) in which no phase shift occurs, capacitance (C_x) in which a 90° phase shift occurs, and the 180° phase reversal between grid and plate allow a theoretical continuously-variable phase shift from 0° to 180°. As the resistance is made less than the capacitive reactance, the phase angle of each stage approaches 0° because the path of the signal favors the resistor through which no phase shift occurs. As the resistance is made to exceed the capacitive reactance, the phase angle of each stage approaches 180° because the path of the signal favors the capacitor through which a phase shift occurs. Theoretically each stage should provide a phase shift from 0° to 180° as the resistance (R_3) is varied. By actual measurement each stage provides a shift of approximately 130°. For this reason three stages are necessary to provide a continuous 360° phase shift. The fourth stage is used as a buffer amplifier to prevent following circuitry from loading the phase-shifting circuit.

Construction Details

The unit can be constructed in accordance with the schematic of Figure 2. The value of the plate and cathode resistors (R_1) is not critical yet they should be matched for each stage. The three phase-shift potentiometers (R_3), ganged on a common shaft, are linear with a minimum power rating of two watts. Although a Heliopot potentiometer, Model 5703, is incorporated in the unit under discussion, any similar potentiometer will suffice.

With regard to the value of C_x, three capacitors (.03, .01, and .005 μf) are mounted on each section of a three-section three-position rotary switch. Each capacitive value allows continuous phase shift over a particular limited

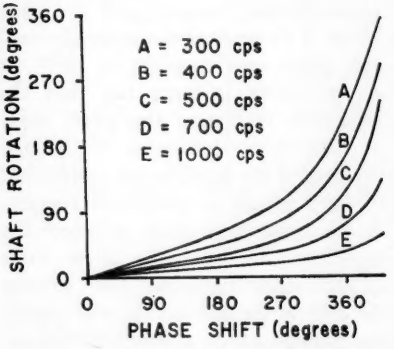

FIGURE 3. Family of curves showing degree of phase shift as a function of phase-shift potentiometer (R_s) shaft rotation with the .01 μf capacitor in the circuit. Frequency is the parameter. Capacitive reactance (X_c) equals resistance (R) at 400 cps.

frequency range without excessive amplitude fluctuation. Alternate switching of the three capacitive values into the circuit enables phase shift to be effected from 150 to 8000 cps.

Circuit Characteristics

At frequencies where capacitive reactance is equal to resistance, amplitude variation is less than plus or minus 0.5 db as phase is shifted through 360°; at frequencies one octave removed, the variation is essentially plus or minus 1.5 db. Harmonic analysis disclosed that the second harmonic is more than 30 db below the fundamental at all frequencies regardless of phase shift.

Degree of phase shift as a function of phase-shift potentiometer shaft rotation is shown in Figure 3 with frequency as the parameter. The family of curves shows 360° phase shift from 300 to 1000 cps with the .01 μf capacitor (C_x) in the circuit. Curves with the .03 and .005 μf capacitors in the circuit disclosed similar results for lower and higher frequency ranges, respectively. All phase values were recorded from a phase angle meter (Acton, Model 350AB).

Some workers may view as a major shortcoming the fact that as the phase-shift potentiometer shaft is rotated, phase shift is not linear through 360°. This would present a problem only for those who do not have phase-monitoring equipment at their disposal.

Summary

The purpose was to describe a method of shifting audiovoltage phases continuously through 360°. The unit in which the method is applied is both simple and economical to construct. Construction details and circuit characteristics are discussed.

References

1. Bureau of Ships, Navy Department, *Radar Electronic Fundamentals*, Navships 900,-014. Washington: U. S. Gov. Ptg. Off., 1944.
2. CHANCE, B., and others (Eds.), *Electronic Time Measurements*, Mass. Inst. Tech., Radiation Laboratory Series, vol. 20. New York: McGraw-Hill, 1949.
3. DOLCH, J. P., An investigation of some phase and intensity relationships in the interference of bone- and air-conducted sound. Univ. Iowa, Signal Corps; Cont. DA-36-039 sc-42562, *Supplementary Report* 2, 1953.
4. EVEREST, F. A., Phase shifting up to 360 degrees. *Electronics*, 14, 1941, 46-49, 122.
5. MAHREN, A. A., Synthetic waveforms speed wave analysis. *Electronics*, 25, 1952, 132-135.

Electrodermal Response Audiometry with Mentally Defective Children

JAMES W. MOSS

MARGARET MOSS

JACK TIZARD

The primary purpose of the present study was to evaluate the utility of electrodermal response (EDR) audiometry with mentally defective children. Because conditioned audiometry utilizing EDR measures does not require the active participation of the subjects tested, it has been suggested (*6*) as a possible method of assessing the hearing of mentally deficient children. The experiences of the present authors suggest that the utility of such a technique with defectives may be open to question.

Irwin, Hind, and Aronson (*4*) working with 20 institutionalized mentally defective children (median IQ 50; age range 9 to 38 years), found that EDR audiometry was not as effective as standard audiometry. They evaluated the hearing of their subjects using both EDR and standard methods and reported that in only half of their cases were EDR thresholds within 10 db of those established by standard methods. A quarter of their subjects were untestable by EDR procedures because of inadequate conditioned responses or because of behavioral problems. Kodman, Fein, and Mixson (*5*) reported that EDR audiometry gave 'important information concerning auditory sensitivity' in 28 out of 31 cases. Three of their subjects were rejected as unmanageable. The authors used a classical conditioning procedure with an interval of .5 sec between the unconditioned stimulus and the conditioned stimulus. In neither of the above studies is there any discussion of the selection of subjects except that they were described as mentally defective residents of state institutions.

Difficulties in the use of conditioned EDR audiometry have been reported by some investigators working with nonretarded children. Goldstein (*2*), working with subjects suspected of having nonorganic hearing losses, found that about half of the group gave unreliable EDR recordings. He concluded that the EDR method fails sufficiently

James W. Moss (Ph.D., George Peabody College, 1958), Postdoctoral Fellow, Institute of Psychiatry, Maudsley Hospital, London, is now at the University of Illinois. Margaret Moss (A.B., Vanderbilt University, 1957) was attached to the Psychology Department, Institute of Psychiatry, Maudsley Hospital, London, while this study was in progress. Jack Tizard (Ph.D., University of London, 1951) is with the Medical Research Council, Social Psychiatry Research Unit, Institute of Psychiatry, Maudsley Hospital, London. This project was supported in part by Research Grant MF 8599 from the National Institute of Mental Health, Bethesda, Maryland.

often that lack of response to tone could not be equated with loss of hearing. Goldstein, Polito-Castro, and Daniels (*3*) found that 14 out of 32 normally-hearing children were difficult to condition with EDR techniques.

Studies dealing with EDR audiometry frequently emphasize the clinical utility of the method without regard to such factors as the percentage of reinforced trials or the existence of an unconditioned response to the tone. This may be justifiable in the clinical situation where the primary concern is the elicitation of a response to the tone. It is frequently possible, however, to elicit a response to a loud tone where no conditioning has taken place; but without conditioning a child who responds to a relatively loud tone may not give a response to a near-threshold tone, thereby creating the false impression of a hearing loss. As one of the interests of the present study was the conditioning process itself, an attempt was made to control for the presence of an unconditioned response to tone by adaptation trials prior to the test series.

Specifically, the present study had three aims. These were (a) to determine the extent to which mentally defective children could be conditioned with EDR methods, (b) to examine the reliability with which the records could be interpreted, and (c) to describe the management problems associated with testing.

An attempt was made also to see whether a bright light could be as effective as a mild electric skin stimulus for the unconditioned stimulus. This was considered of value because some patients are unable to tolerate having electrodes placed on them, and an alternative unconditioned stimulus would be useful.

Method

Subjects. The design of this study originally called for a random selection of 24 institutionalized mentally defective subjects who were ambulatory and between six and eight years in chronological age. A random selection of patients was wanted so that generalizations could be made concerning the utility of the EDR methods with institutional defectives in general rather than with a selected group of defectives. The relatively low age level was selected in order that information could be obtained on children who would be most likely to have difficulties with standard audiometric methods.

A list of patients meeting these criteria was drawn from the hospital records. The first selections from this group, however, were found to be so severely unmanageable that their participation in the study was out of the question. It became evident that the majority of children selected in this random fashion could not be expected to cooperate even to the limited extent required for EDR conditioning.

It was accordingly decided to select the children by asking the teachers in the school program for the names of their most cooperative children. Even with this selection system, 34 children had to be tested before 24 complete records were obtained. The final sample upon which completed records were obtained consisted of 24 subjects equally divided as to sex. Of these, 13 were mongols. The chronological ages ranged from five years, four months, to

15 years, two months, with a mean of nine years, 10 months, and a standard deviation of 28 months. There were no significant differences between the sexes in chronological age. The median IQ for the group was 37. The Stanford-Binet scale had been used with 18 of the children. The remaining scores were obtained by use of the Merrill-Palmer, the Vineland, or the Griffiths scales. All but two of the children had been tested within the two-year period immediately prior to the study.

With the exception of the mongols, the group was heterogeneous as regards diagnosis and aetiology. One subject had cerebral palsy, two were postinfectious cases, in two there was a history of a maternal toxemia, one had tuberous sclerosis, and the remainder were of unknown or undifferentiated aetiology.

The 10 children of the 34 tested who did not produce complete records were generally rejected for behavioral reasons such as negativism, hyperactivity, shyness, or withdrawal. One child had a severe cough which interfered with the record. Another was quite cooperative but her record, after a 45 minute run, never settled sufficiently to permit the conditioning trials to start. The mean CA of this rejected group was eight years, one month and the median IQ was 24.

Apparatus. A two-channel AC EEG amplifier was used to measure skin potential which was recorded on an Ediswan pen-oscillograph. The paper speed was 1.5 cm per sec. A Peters audiometer was used to produce the conditioned stimulus.

Two pickup electrodes were used, one on the inner surface of the left wrist, the other on the palmar surface of the proximal phalanx of the left middle finger. These electrodes were silver discs 8 mm in diameter, resin mounted in a rubber ring 15 mm in diameter allowing for a 1 mm space for electrode jelly. The ground electrode, placed on the inner surface of the left forearm, was a 19 mm silver rectangle with a raised plastic border again allowing space for jelly. The skin stimulus electrodes (17 mm x 20 mm, silver) were attached to the calf of the right leg. All electrodes were secured with adhesive tape. Cambridge electrode jelly was used to insure maximum conductibility.

A 1000 cps, 65 db pure tone was produced by the audiometer and fed binaurally to the subjects through earphones. The skin stimulus was derived from 50 cps alternating current mains through an isolating transformer. In order to minimize mains pickup by the amplifiers, the source was balanced to a ground. The electrodes were each connected through a 4.7 meg resistor to a twin-gang potentiometer by which the experimenter could vary the voltage from 0 to 500 v. The maximum current that could be passed through the subject was limited to just over 0.5 ma.

Three electric photoflood lamps (240/250 v, 275 w) served as the alternative unconditioned stimulus. These were placed approximately two feet in front of the subject at eye level in a row across his field of vision. Only one of the three lights was flashed at a time in a prearranged pattern.

Procedure. The timing within each presentation was automatically controlled. The experimenter pressed a

key which started a tone of 1.6 sec duration. The unconditioned stimulus started 0.6 sec after the onset of tone and the two terminated simultaneously. All equipment was screened from the subject's view.

Each subject was individually tested on two occasions, once using the skin stimulus and once using light as the unconditioned stimulus. Half of the group received the skin stimulus first; the remainder received the light condition first.

The subject was seated before a table on which were placed the three lamps and a number of magazines and pictures. He was made as comfortable as possible with the leads and headphones in place. With some subjects it was necessary to restrain the left arm and hand to prevent gross movement.

The testing consisted of five parts.

a. The 1000 cps tone of 65 db was presented at approximately 20 sec intervals until any unconditioned response to the tone was extinguished or adapted. Three consecutive tone presentations without response were accepted as the criterion that the tone was not still serving as an unconditioned stimulus. A 65 db tone level was used throughout to insure that the child could hear the tone and to maximize the chances of obtaining clear-cut results.

b. The unconditioned stimulus (light or skin stimulus) was presented. Three trials were given with the lights. Sufficient trials were given with the skin stimulus to determine the starting stimulus magnitude. The level of stimulation was considered adequate if an EDR was elicited on three successive trials.

c. The tone was again presented alone for three or more trials as a final check to be sure that it was not serving as an unconditioned stimulus. Where a response occurred to the tone, sufficient trials were given for adaptation to a criterion of three consecutive presentations without a response.

d. The tone was presented for 48 trials at approximately 20 sec intervals. For 30 of these trials the unconditioned stimulus was paired with the tone. For the remaining 18 trials, which were randomly interspersed among 30 conditioning trials, the tone alone was presented. These served as test trials of whether conditioning had in fact taken place. The order of presentation of conditioning and of test trials was the same for each subject. When a subject failed to give an EDR to a skin stimulus trial the intensity of the stimulus was increased.

e. Additional tone-alone trials were presented in order to obtain a measure of extinction. These trials were continued until no responses were recorded for three consecutive trials.

Results

Complete records were obtained from 24 subjects under both the light and skin stimulus conditions. Each complete record was classified as *good*, *fair*, or *poor* by two raters, the terms being defined as follows:

Good, record stable with a minimum of spontaneous fluctuations, responses generally scorable with confidence;

Fair, labile record with spontaneous fluctuations but responses still generally scorable;

Poor, many spontaneous fluctuations, or inadequate unconditioned responses, scoring difficult or impossible.

Only 13 of the 48 records could be rated as *good* or *fair* by the two raters, four of the 13 being obtained with the lights and nine with the skin stimulus. The remaining 35 records were difficult to score with confidence and, in some cases, were completely unscorable.

Individual responses were next rated according to the ease of scoring and the degree of confidence which the judges placed in their scoring. For the 13 *fair* and *good* records, 77% of the responses on the reinforced trials were considered as well defined, clear-cut, positive responses which could be easily scored. Rater agreement was 87%. Attempts were made to score the responses on records rated as *poor*, but the reliability of the ratings was low. Only 32% of the responses to the reinforced trials were considered as clearly positive by both raters. The rest were either impossible to score with confidence or there was no response at all. The agreement between raters for this group was 65%.

When only conditioned responses were considered, the percentage of scorable, positive responses dropped to 32% for the *fair* and *good* records and to 11% for the rest.

It was found that 95% of the subjects gave unconditioned responses to the tone alone during the pretrials. The number of responses during the pretrials was correlated (r=0.42) with the number of conditioned responses during the actual conditioning trials ($p<0.05$). Because of the difficulties in the scoring and the unreliability inherent in the records this finding must be considered as only suggestive.

No attempt was made to test the hearing of the 24 subjects with standard methods, so it is not known to what extent this was possible. It was later found, however, that two of the children had been tested by standard audiometric techniques. The audiometrist reported that in both cases hearing was normal and the tests had been carried out without difficulty.

Many of the children were difficult to control during testing and they became progressively more so towards the end of the test. One child, for example, became highly excited when the light flashed. He would strike out at the lamp and bang the table. On the nonreinforced trials, he would strike at the lamp which lit the preceding time. Other children simply became more and more restless and active.

A few of the children, however, were quite passive and lethargic. One actually went to sleep during the trials. Her record smoothed out and showed good conditioned responses during the time that she was asleep, though when she was awake her record was difficult to interpret. Others simply sat in their place without moving for the entire period.

Because of the poor quality of most of the records, it was decided to check the stability of the equipment by testing three normal children. The records of the normal children were easily scored with high agreement between the judges, suggesting that the equipment was in good working order.

Discussion

Both the evaluation of conditioning in severely defective children, and the study of bright lights as an alternative unconditioned stimulus, had to be

abandoned because of the unreliability of the records. Inasmuch as the records of the three normal children were essentially normal and could be reliably scored, it was concluded that the unreliability of the defectives' records was a function of their deficiency and not an artifact of the method or equipment.

The data from the present study suggest that only a small percentage of children with severe mental defects yield reliable and useful EDR records under hearing-test conditions. Not only were all of the first subjects selected on a random basis found to be untestable, but even with a group of the more carefully selected subjects, 10 out of 34 failed to produce complete records. Furthermore, only 13 of the 48 records finally obtained allowed interpretation with any degree of accuracy and reliability.

These findings raise a question concerning the utility of the EDR method for testing the hearing of young children with severe mental defects. The reports from other investigators do not indicate how their subjects were selected, but one wonders if their subjects were representative of institutionalized children with severe mental defects as a whole, or if only the most cooperative children were tested. Kodman, Fein, and Mixson (*5*) discussed the need to work with children who could not be tested by standard methods but they do not indicate whether their subjects had, in fact, failed to cooperate in such procedures. In the study reported by Irwin, Hind, and Aronson *(4)*, it is known that most of the subjects were capable of being tested by standard audiometric procedures, only three out of 20 subjects having failed to cooperate sufficiently for standard methods to be used while five of the 20 could not be tested by EDR methods.

The implications of the above discussion are consistent with the conclusions drawn by Irwin, Hind, and Aronson, that is, that there is some doubt as to whether the unpleasantness and complications of EDR testing can be justified with subjects who can be tested in other ways. The present experimenters go one step further, however, and suggest that where a mentally defective subject is unable to participate in normal audiometric operations, he is probably unable also to participate successfully in EDR audiometry.

The finding that most of the children in this study gave an unconditioned response to the tone alone and that the number of such responses was correlated with the number of conditioned responses given during the test series suggests that some of the children may not have been conditioned at all, but rather, that they were simply responding to the tone as well as to the skin stimulus or light. Generally speaking, the responses to the tone, after adaptation during the first pretrial, appeared again for the second pretrial. There is reason to believe that these unconditioned responses elicited during the pretrial series appeared again during the test series as 'unconditioned' responses. This finding points out the need for careful attention to the conditioning procedures in EDR audiometry. The underlying assumption in EDR audiometry is that the subject will associate

the tone with the shock and that, because of this association, he will give a physiological response on hearing the tone. Where the response appears for any reason other than the association established between the tone and the shock the clinician is not in a position to interpret the record. If a loud tone can serve as an unconditioned stimulus, and if EDR is related to the magnitude of the unconditioned stimulus as Aronson, Hind, and Irwin *(1)* suggest, then one would expect smaller and fewer responses at a near-threshold level than at a loud level. The clinician who assumes that any EDR deflection following tone is adequate for the testing of hearing may find himself diagnosing hearing losses where none exist.

Summary

The purpose of this study was to examine conditioned electrodermal response (EDR) techniques of audiometry with mentally deficient children and to evaluate conditioning under two stimulus situations. The 24 subjects were tested once with a mild electric skin stimulus and once with a bright light as the unconditioned stimulus.

The EDR records as a whole were rated as *good*, *fair*, or *poor* by two judges. Individual responses were rated according to the ease of scoring and the degree of confidence which the judges placed in their ratings. Only 13 of a total of 48 records were rated as *good* or *fair*.

It is concluded that EDR audiometry with mentally defective children is of limited utility. It seems likely that most mentally defective children who cannot be tested by standard pure-tone methods probably cannot be tested by EDR methods.

Acknowledgment

The authors wish to express their thanks to Dr. B. H. Kirman, Consultant Psychiatrist, the Fountain Hospital, London, the Drs. P. H. Venables and N. O'Connor of the Medical Research Council, Social Psychiatry Research Unit, and to Miss B. A. Levy, speech therapist at the Fountain Hospital, for their cooperation in carrying out the study.

References

1. Aronson, A. E., Hind, J. E., and Irwin, J. V., GSR auditory threshold mechanisms: effect of tonal intensity on amplitude and latency under two tone-shock intervals. *J. Speech Hearing Res.*, 1, 1958, 211-219.
2. Goldstein, R., Effectiveness of conditioned electrodermal responses (EDR) in measuring pure-tone thresholds in cases of non-organic hearing loss. *Laryngoscope*, 66, 1956, 119-130.
3. Goldstein, R., Polito-Castro, Sara B., and Daniels, Jane T., Difficulty in conditioning electrodermal responses to tone in normally hearing children. *J. Speech Hearing Dis.*, 20, 1955, 26-34.
4. Irwin, J. V., Hind, J. E., and Aronson, A. E., Experience with conditioned GSR audiometry in a group of mentally deficient individuals. *Train. Sch. Bull.*, 54, 1957, 26-31.
5. Kodman, F., Jr., Fein, A., and Mixson, A., Psychogalvanic skin response audiometry with severe mentally retarded children. *Amer. J. ment. Defic.*, 64, 1959, 131-136.
6. Schlanger, B. B., and Gottsleben, R. H., Testing the hearing of the mentally retarded. *Train. Sch. Bull.*, 54, 1957, 21-25.

An Articulation Test for Assessing Competency of Velopharyngeal Closure

HUGHLETT L. MORRIS

D. C. SPRIESTERSBACH

FREDERIC L. DARLEY

It has been established that the competency of velopharyngeal closure is directly related to the speech problems of individuals with cleft palates. Two techniques are commonly used for inferring the competency of closure during speech: (a) Lateral head x rays may be taken during a speech production which can be sustained and which is assumed to require maximum voluntary closure; the films are inspected to determine whether the soft palate approximates the posterior pharyngeal wall or adenoidal tissue (*3, 9, 11, 16*). (b) Oral pressure ratios may be obtained through the use of any technique which enables one to measure the magnitude of the pressure in the oral cavity with the nostrils open and occluded (*4, 9*).

It is the thesis of this paper that results of articulation testing may be used as another index of the adequacy of velopharyngeal closure. Spriestersbach, Darley, and Rouse (*8*), Counihan (*3*), Starr (*10*), Bzoch (*2*), Subtelny and Subtelny (*12*), and McWilliams (*6*) have reported that cleft palate speakers typically have most difficulty with plosive and fricative sounds. This finding indicates that performance on articulation test items which include sounds requiring intraoral pressure may reflect adequacy of velopharyngeal closure for purposes of speech production. Although the investigators above do not report specific information regarding the adequacy of velopharyngeal closure of the subjects studied, it may be inferred from their descriptions that the majority of subjects did not have competent closure.

The purpose of the present research was to construct an articulation test (hereinafter referred to as the Iowa Pressure Articulation Test) for assessing the adequacy of oral pressure for speech sound production and thus, inferentially, the adequacy of velo-

Hughlett L. Morris (Ph.D., University of Iowa, 1960) is a Research Associate, Department of Otolaryngology and Maxillofacial Surgery, University of Iowa. D. C. Spriestersbach (Ph.D., University of Iowa, 1948) is Professor, Department of Otolaryngology and Maxillofacial Surgery and Department of Speech Pathology and Audiology, University of Iowa. Frederic L. Darley (Ph.D., University of Iowa, 1950) is Associate Professor, Department of Speech Pathology and Audiology, University of Iowa. This investigation was supported in part by research grants M-1158 and D-853 from the National Institutes of Health, Public Health Service.

TABLE 1. Descriptive data concerning subjects comprising the good and poor closure groups (N=25 in each group). Items are age in years and months; sex; cleft type, (a) lip and palate and (b) palate only; full-scale IQ; socioeconomic status, upper, middle, lower; wet spirometer ratios; x-ray readings, (a) /u/ closure complete and (b) /s/ closure complete.

Item	*Statistic*	*Good Closure Group*	*Poor Closure Group*
	Range	5:3-15:3	5:1-15:5
Age	Mean	10:3	10:5
	SD	3:2	3:5
Male	N	18	18
Female	N	7	7
Lip and Palate	N	22	22
Palate Only	N	3	3
	Range	75-128	62-125
IQ	Mean	97.72	96.12
	SD	13.35	15.25
Upper	N	0	1
Middle	N	14	15
Lower	N	11	9
	Range	.94-1.0	.11-.83
Ratio	Mean	.99	.47
	SD	.03	.21
/u/ Closure	N	14	0
/s/ Closure	N	22	1

pharyngeal closure. Specifically, an answer to the following question was sought. Which articulation test items discriminate best between two groups of speakers with cleft palates, one group with adequate and one with inadequate velopharyngeal closure?

Procedure

Subjects. Two groups of 25 subjects each were selected. All subjects were Caucasian singletons (not twins) with medical diagnoses of cleft lip and palate or cleft palate only; no subjects with cleft lip only were included. All cleft palates had been closed surgically; there were no unmanaged or prosthetically managed cleft palates. The 25 subjects in the good closure group had complete closure on x-ray films while sustaining either the /s/ or /u/ sound and a wet spirometer ratio[1] of .90 or higher. The 25 subjects in the poor closure group could not meet the criteria for inclusion in the good closure group and were selected to match those in the good closure group for sex, age, intelligence, and socioeconomic status.

[1]The wet spirometer ratio is obtained by dividing the reading obtained with nostrils open by the reading obtained with nostrils occluded.

TABLE 2. The 74 test items on which there was a significant between-group difference in number of errors. Items are rank-ordered according to magnitude of these differences.

Group Difference	*Single Items*	*Two-element Items*	*Three-element Items*
17		/-sl̩/	
16	/s-/	/sk-/	
15	/-k-, z-/	/ʃr-, -lz, sm-, -sm, sn-/	/-stɚ, -skɚ, -ɚst, spr-, str-, -ntθ/
14	/-z-, -z, ʃ-/	/-kɚ, st-, -st, sw-/	/-skw/
13	/-s-, -ʃ-/	/kr-, -lθ, sl-, -zm, -dʒd, -kt/	/-lfθ, spl-, skr-, -kst/
12	/-ŋ-, -s/	/tr-, gr-, -gɚ, -ɚk, -sk, -pt, -ɚg, kl-, gl-, -gl̩, -lk, sp-, -sp/	/-ŋkl̩, -mps/
11	/k-, g-, -g, -ʃ, dʒ-, -dʒ/	/-ʃɚ, -ɚdʒ, bl-, -ks/	
10	/-k/	/br-, dr-, tw-/	/-ŋgl̩/
9	/t-, -f-, -f/	/-pɚ, pl-, -lf/	

Table 1 indicates that, in general, the two groups were similar in age, sex, Full-Scale IQ score,[2] and family socioeconomic status, as estimated through use of the Index of Status Characteristics (*15*). Group differences in mean wet spirometer ratios and in velopharyngeal closure as indicated by x-ray films were felt to be of sufficient magnitude for the purposes of this study.

The oral mechanisms of the subjects were judged with regard to the following characteristics selected for their importance in speech production: (a) continuity of cutting edge of anterior teeth, (b) palatal contour, and (c) anteroposterior relationship of upper and lower incisors. The ratings were not made, however, with reference to the effect of these rated variables on speech, nor with reference to the normalcy of either the structures or of the structural relationships of the parts of the oral mechanism. The task was simply one of description. Dental models for the 50 subjects were presented in random order to three trained judges, who rated them by mutual agreement. No measurements were made; spatial relationships and dimensions were judged. No important difference between the two groups was obtained for any of the judgments of the three selected characteristics.

In addition, measurements of general facial morphology were obtained from lateral head x-ray films taken with the teeth in occlusion. Measurements were made of angulation of cranial base, protrusion or retrusion of maxillary incisors in relationship to the facial plane, and the anteroposterior relationship of the maxilla and mandible. No important group differences were found regarding these measurements.

Since the two groups of subjects were matched with respect to factors considered important in the development of articulation skills for normal

[2]The Wechsler Intelligence Scale for Children was administered to all subjects.

children (age, sex, intelligence, and socioeconomic status) and since no differences were found with regard to general physical traits which may be important in speech production (type of cleft, type of management, description of oral mechanism, and general facial morphology), it was assumed that any variation in articulatory proficiency that existed could best be accounted for by differences in ability to impound air in the oral cavity and thus ability to achieve closure.

Testing Procedure. The 176-item Templin-Darley Diagnostic Test of Articulation (*14*), using pictorial presentation to elicit spontaneous single-word responses, was administered to all subjects. The test was scored for the number of items produced correctly.

Results

General Results. The total number of errors per closure group was obtained for each item on the 176-item test. The items were then rank-ordered according to the magnitude of the differences between the number of errors made by the two groups. Group differences in number of errors were very small for vowel and diphthong items,

TABLE 3. Single-consonant items on which there were significant differences between the number of errors made by the two closure groups classified according to manner of articulation. Items are rank-ordered according to magnitude of intergroup differences.

Group Difference	*Single-Consonant Items*				
	Fricatives	*Plosives*	*Affricates*	*Nasals*	*Glides*
16	/s-/				
15	/z-/	/-k-/			
14	/-z-, -z, ʃ-/				
13	/-s-, -ʃ-/				
12	/-s/	/-g-/			
11	/-ʃ/	/k-, g-, -g/	/dʒ-, -dʒ/		
10	/-ʒ-/	/-k/	/-dʒ-/		
9	/f-, -f, -f-/	/t-/			
8	/-θ/	/-t/	/-tʃ-/		
7	/v-, -v, -ð, -ʒ/	/-p-, d-/	/-tʃ/		
6	/-v-, θ, -θ-, ð-/	/-p/	/tʃ-/	/-ŋ/	
5		/-t-, -d-, -d/			/r-/
4		/p-/			/-r-, l-/
3	/-ð/	/b-, -b/		/-n-/	/-l-, -m-, m-, j-, -j-/
2		/b-/		/n-, -n, -ŋ-/	
1					/-l, h-, -h-/
0				/m-, -m-, -m/	/w-, -w-/

ranging from eight errors for /ɝ/ to no difference for /ʌ/, /ɑ/, /ɔ/ and /ɔɪ/. The distributions of the group differences for singles and for two-element items have wide ranges, 17 and 19 points, respectively. The distribution of the group differences for the three-element items has a range of 16 points. The latter distribution is extremely skewed, however; 13 of the 16 three-element items are distributed over a four-point range.

Differences between the number of errors made by the two groups were evaluated by the chi square test of independence. For 74 of the test items (Table 2) these differences were significant ($\chi^2_{.01} = 6.6$, $df = 1$). These 74 items included one of the 18 vowels, 19 of the 68 consonant singles, 40 of the 74 two-element items, and 14 of the 16 three-element items. The single vowel was /ɝ/.

The greatest between-group error differences on the consonant singles of the articulation test were in the fricatives and plosives (Table 3); next greatest were in the affricates, then the nasals and glides. Error totals on consonant blends were similar for the two groups.

Position of the sound elements in the test words appears to have no differential discriminatory power between good and poor closure; sound elements in the initial position of the test words are distributed in discriminating power in essentially the same manner as are elements in the medial or final position. In addition, for two- and three-element items there is no essential difference in discriminatory power for syllabic position in the test word.

Differences According to Number of Elements. Spriestersbach, Darley, and Rouse (*8*) have reported that cleft palate speakers have more difficulty with consonant blends than with consonant singles. In their study, subjects, regardless of efficiency of velopharyngeal closure, were tested on 69 singles, 74 two-element items, and 16 three-element items. However gross they may be, the results suggest that there is a differential response according to the number of elements in the test item.

In the present study 33 items were chosen (Table 4) so that the foregoing finding could be examined. Each of 11 consonants occurred in three types of test item: as a single, in a blend with one other element, and in a blend with two other elements. In the scoring of the blend items, the production of only the specified consonant was considered. For example, in the case of /-s/, the correctness of the production of only /s/ was considered in the /-st/ and

TABLE 4. Items selected for study of performance by cleft palate subjects on articulation test items containing different numbers of sound elements.

Single Consonants	*Two-element Blends*	*Three-element Blends*
/s-/	/sp-/	/spr-/
/-s/	/-st/	/-stɝ/
/k-/	/sk-/	/skw-/
/-k/	/-sk/	/-skɝ/
/-g/	/-gl/	/-ggl/
/t-/	/st-/	/str-/
/-t/	/-kt/	/-kst/
/p-/	/pl-/	/spl-/
/-p/	/-pt/	/-mpt/
/-b/	/-bɝ/	/-mbɝ/
/-f/	/-lf/	/-lfθ/

TABLE 5. Summary table for comparison of performance by good (G) and poor (P) closure groups on articulation test items containing increasing numbers of sound elements.

Statistic	*Single Items*		*Double Items*		*Triple Items*	
	G	*P*	*G*	*P*	*G*	*P*
Total Number Correct	242	149	230	135	225	97
Mean Score	9.68	5.96	9.2	5.4	9.0	3.88
Variance	1.06	1.68	.50	.97	1.02	1.36
SD	1.03	1.29	.71	.98	1.01	1.17
Median Score	10.5	6.5	11.0	6.5	11.0	3.0
Range of Scores	6-11	0-9	4-11	0-11	2-11	0-9

/-stɔ˞/ items; the correctness of production of the /t/ and /tɔ˞/ was not scored. The number of correct productions of the specified consonant for each of the three types of item by each closure group is presented in Table 5.

Within-Group Differences. The performances of the subjects with adequate closure were similar on all three types of item. The largest differences appear between the variances. The subjects with inadequate closure, on the other hand, performed differently on the three types of item. The number of items produced correctly decreases as the number of consonants in the items increases. Group differences in mean and median scores are greatest for three-element items. The variances differ for the three types of item. The usual analysis of variance technique is thus inadvisable for testing the significance of the differences both because the variances differ greatly and because the distributions of scores are skewed, especially for the poor closure group. A nonparametric technique, the Friedman two-way analysis of variance by ranks procedure (7, pp. 166-172), was employed to test the significance of within-group differences for each item type. The differences between the performances on the three types of test item were not significant for the good closure group ($\chi_r^2 = 1.50$, $df = 2$) but were significant at the 5% level for the poor closure group ($\chi_r^2 = 15.02$, $df = 2$).

Between-Group Differences. The performances by the two groups differ greatly for each type of test item (Table 5). Not only are the mean and median scores different in size, but the poor closure group is more variable in performance than the good closure group for all three types of item. There were increasingly greater between-group differences as the number of elements in the test items increased.

Construction of the Test

The Iowa Pressure Articulation Test was constructed on the basis of the above findings and, beyond this, used only such items as normal young children typically can produce without difficulty. Templin (*13*) reports earliest ages at which 75% of her subjects (who ranged in age from three to eight years) produced the articulation

TABLE 6. Items (selected from the 74-item list) which are reported by Templin to be articulated correctly by 75% of children five years of age. These 43 items constitute the Iowa Pressure Articulation Test.

Group Differences	*Single Items*	*Two-element Items*	*Three-element Items*
16	/s-/	/sk-/	
15	/-k-/	/sm-, -sm, sn-/	/str-/
14	/ʃ-, -z-/	/-kə, st-/	
13	/-s-, -ʃ-/	/kr-/	
12	/-g-, -s/	/sp-, tr-, gr-, -gə, -ək, -pt, kl-, gl-/	/-mps/
11	/k-, g-, -g, -ʃ, dʒ-/	/-ʃə, bl-, -ks/	
10	/-k/	/br-, dr-, tw-/	
9	/t-, -f-, -f/	/-pə, pl-, -lf/	

test items correctly. Presented in Table 6 are the 43 items, out of the present study's total of 74 items that yielded significant between-group differences, that are reported by Templin (*13*) to be correctly produced by 75% of children five years of age. Of these, 17 are single consonant items, 24 are two-element items, and two are three-element items. These 43 items constitute the final selection of items for the Iowa Pressure Articulation Test, designed to assess the adequacy of velopharyngeal closure of cleft palate speakers.

Discussion

The results of this study are generally consistent with other research findings. Fricatives, plosives, and affricates prove to be the best discriminators between speakers with adequate and speakers with inadequate velopharyngeal closure. Hudgins and Stetson (*5*) and Black (*1*) have shown that oral breath pressure during the production of these phoneme types is greater than during the production of any other types. One would expect performances on these items to reflect differences in ability to impound such pressure and, inferentially, ability to achieve velopharyngeal closure.

Fricatives show the greatest between-group differences. Subtelny and Subtelny (*12*) report that fricatives are the most sensitive to inadequate velopharyngeal closure while plosives may be correctly produced with small velopharyngeal openings, as shown by measurements on x-ray film. The x-ray film criterion used for the selection of subjects in this study considered only whether closure was complete or incomplete. The possibility exists that there was a variation in the degree of closure within the poor closure group. This possibility is supported by the range of the obtained wet spirometer ratios for the poor closure group (.11 to .83). To the degree that variation in wet spirometer ratios reflects variation in the size of the velopharyngeal opening, one might expect a variety of differences among individuals to be

obtained for the plosive items, which differences would generally be smaller than those obtained for the fricative items.

Summary

The 176-item Templin-Darley Diagnostic Test of Articulation was administered to 50 children with cleft palates, one group of 25 with adequate and one group of 25 with inadequate velopharyngeal closure. On the basis of obtained group differences the 43-item Iowa Pressure Articulation Test was constructed for assessing adequacy of velopharyngeal closure in cleft palate speakers.

The following conclusions seem warranted: (a) Articulation test items containing fricatives, plosives, and affricates are the best discriminators between speakers with adequate and inadequate velopharyngeal closure. (b) Position of sound element in the test word apparently has no differential discriminatory effect on the articulation performances of individuals with adequate and inadequate closure. (c) Cleft palate speakers with adequate velopharyngeal closure give essentially similar performances on consonant singles, two-element items, and three-element items, while cleft palate speakers with inadequate velopharyngeal closure have increasingly greater difficulty as the number of elements in the test items is increased.

References

1. Black, J. W., The pressure component in the production of consonants. *J. Speech Hearing Dis.*, 15, 1950, 207-210.
2. Bzoch, K. R., An investigation of the speech of pre-school cleft palate children. Ph.D. dissertation, Northwestern Univ., 1956.
3. Counihan, D. R., A clinical study of the speech efficiency and structural adequacy of operated adolescent and adult cleft palate persons. Ph.D. dissertation, Northwestern Univ., 1956.
4. Goddard, Lisette, A normative study of oral breath pressure in children. M.A. thesis, Univ. Iowa, 1959.
5. Hudgins, C. V., and Stetson, R. H., Voicing of consonants by depression of larynx. *Arch. neerl. Phon. exp.*, 11, 1935, 1-28.
6. McWilliams, Betty Jane, Articulation problems of a group of cleft palate adults. *J. Speech Hearing Res.*, 1, 1958, 68-74.
7. Siegel, S., *Nonparametric Statistics*. New York: McGraw-Hill, 1956.
8. Spriestersbach, D. C., Darley, F. L., and Rouse, Verna, Articulation of a group of children with cleft lips and palates. *J. Speech Hearing Dis.*, 21, 1956, 436-445.
9. Spriestersbach, D. C., and Powers, G. R., Articulation skills, velopharyngeal closure, and oral breath pressure of children with cleft palates. *J. Speech Hearing Res.*, 2, 1959, 318-325.
10. Starr, C. B., A study of some characteristics of the speech and speech mechanisms of a group of cleft palate children. Ph.D. dissertation, Northwestern Univ., 1956.
11. Subtelny, Joanne D., A laminagraphic study of nasalized vowels produced by cleft palate speakers. Ph.D. dissertation, Northwestern Univ., 1956.
12. Subtelny, Joanne D., and Subtelny, J. D., Intelligibility and associated physiological factors of cleft palate speakers. *J. Speech Hearing Res.*, 2, 1959, 353-360.
13. Templin, Mildred C., Certain language skills in children, their development and interrelationships. *Inst. Child Welf. Monogr.*, Ser. 26. Minneapolis: Univ. Minnesota Press, 1957.
14. Templin, Mildred C., and Darley, F. L., *The Templin-Darley Tests of Articulation*. Iowa City: Bur. Ed. Res. and Serv., Ext. Div., Univ. Iowa, 1960.
15. Warner, W. L., Meeker, Marchia, and Eells, K., *Social Class in America*. Chicago: Science Res. Assoc., 1949.
16. Wildman, A. J., The role of the soft palate in cleft palate speech. *Angle Orthodont.*, 28, 1958, 79-86.

Pitch Discrimination and Articulation

RONALD K. SOMMERS

WILLIAM J. MEYER

ANN K. FENTON

Auditory perceptual behavior and hearing acuity have been causally linked by various investigators (*2, 11, 18*) to voice and articulation problems of children. Some researchers have found a significant relationship between auditory discrimination skills and defective articulation. Although normative data have been gathered by Templin (*15*) for auditory discrimination ability of children of young ages, the factors that comprise auditory discrimination have been largely unstudied.

In efforts to discover important variables in auditory perceptual behavior which are related to speech disorders, a number of investigators have applied measures of musical talent to individuals with speech problems. In an early study Travis and Davis (*17*) applied the Seashore test of pitch and tonal memory to a large sample of college students and found that speakers with misarticulations on the average have significantly poorer pitch discrimination than do normal speakers. Gilkinson (7), using the Seashore tests to study the relationship between judges' ratings of variables in the speakers' voices and measures of musical talent, found that there was little relationship between speech skill and Seashore scores when general criteria of speech were used. In a similarly designed study, Horowitz[1] found a slight relationship between speech skills of normal speakers and the scores of rhythm and pitch discrimination on the Seashore test. In other studies using the Seashore tests, Eisenson, Kastein, and Schneiderman (*4*) found that subjects with voice problems had significantly poorer pitch discrimination than matched normals. Mange (*12*), using 35 pairs of elementary school age children matched in terms of age, sex, IQ, and classroom placement, reported that those who misarticulated the [r] sound performed less well than normal children on the pitch discrimination subtest.

Bradley (*1*) explored the relationship between pitch discrimination and speech development among three different groups of subjects between the

Ronald K. Sommers (Ed.D., University of Pittsburgh, 1960) is Supervisor of Speech and Hearing, Armstrong County, Pennsylvania. William J. Meyer (Ph.D., Syracuse, 1957) is Assistant Professor of Psychology and Education, University of Pittsburgh. Ann K. Fenton (B.S., Kent State, 1958) is Speech Therapist, Armstrong County Schools.

[1]L. S. Horowitz, Ability of pitch, intensity, and rhythm discrimination as applied to speech (unpublished study).

ages of nine and 14, all of whom were within the normal range of intelligence. One group consisted of 11 subjects found to be normal in articulation and hearing. A second group of 11 subjects was normal in hearing but had enough problems in articulation to require speech therapy. The third group of 11 subjects had problems in both articulation and hearing. Bradley determined the pitch discrimination abilities of all subjects at the frequencies of 500, 1000, and 2000 cps using a 'just noticeable difference' technique. Both ascending and descending measurements were made at 50 db for each ear. The results of this study indicated that subjects who had normal hearing but problems in articulation were significantly poorer in pitch discrimination than were normal subjects. This investigator concluded that poor pitch discrimination may be a contributing factor in the etiology of functional articulation problems since cues requiring precise discrimination for phonetic analysis may be deficient in children who have these problems.

The present study investigated the possibility that the findings of Mange and of Bradley may be generally applicable to children in grades three through 12 who have articulation problems. Furthermore, the attempt was made to determine whether or not significant differences in pitch discrimination exist between groups of school children having articulation errors on [r] and [s], two sounds commonly misarticulated. This seemed important since Mange's findings were based on misarticulation of [r] only, and Bradley did not indicate the types of articulation errors of his 11 normal hearing subjects who had speech problems.

Procedure

The experimental group consisted of 65 school children with speech problems, drawn from the therapy rolls of Armstrong County, Pennsylvania, schools, to meet these two criteria: (a) enrollment in the third grade or above, and (b) demonstration of only one defective phoneme, either [r] or [s]. A control group of 65 subjects was drawn from the attendance rolls within the same schools and matched with the experimental subjects on IQ (using the Otis Intelligence test, form B), grade, and sex. In addition, attempts were made to see that each control subject was enrolled in the same classroom or school as was the matching experimental subject. Matching of subjects on IQ was achieved within a plus or minus five IQ point limitation. There were 26 girls and 39 boys in each group.

All subjects were screened for hearing acuity using a Beltone, 10A audiometer. Any subject with a 20 db hearing loss in either ear in the frequency range 250 to 4000 cps was excluded from the study. The articulation of each subject was tested by two of the authors using the Wernock-Medline Picture Articulation Test. The reliability of this articulation testing was determined by having the testers and six other experienced speech therapists test a group of five children with articulation errors. When the articulation scores for each subject were compared for the two experimenters and when these scores were compared for each of the two experimenters and

each speech therapist, the resultant reliability coefficients were above the reported limits (*3, 4, 10*) for acceptability when correct or incorrect judgments are made.

Using standard procedures, each subject was administered the pitch subtest from the Tilson-Gretsch Musical Aptitude Test (*16*). In each case two sample trials were presented so that the subjects understood the task and recorded the appropriate responses on the scoring sheet. The task consisted of listening to 25 pairs of tones and indicating on the scoring blank if the second of each pair was higher or lower than the first. The score for each subject was the total number of correct responses. Generally, three or four subjects were tested simultaneously and, in many instances, these included matched experimental and control subjects.

The test-retest reliability of this pitch subtest was determined by administering it to a fourth grade class of school children. Testing was done with the entire class and repeated one day later.

TABLE 1. Summary of analysis of variance for evaluation of pitch discrimination scores of 65 school children with functional misarticulations and of a matched group of 65 normal children.

Source of Variation	*df*	*ms*	*F*	$F_{.05}$
Between Groups	1	93.00	4.20	4.00
Between Pairs	64	29.31	1.32	1.54
Residual (error)	64	22.15		
Total	129			

The resultant *r* of .86 indicated that the pitch test, as used by the authors, was adequately reliable for research purposes.

Results

An analysis of variance for matched groups, summarized in Table 1, provides evidence to support the conclusion that those children with articulation errors on [r] and [s] perform more poorly on a test of pitch discrimination than do children without articulation errors. The mean number of correct pitch discriminations for the group of subjects with articulation problems was 11.72 with a standard deviation of 5.16; for the group of normal subjects the mean was 13.85 with a standard deviation of 5.75.

TABLE 2. Summary of analysis of variance for evaluation of pitch discrimination scores of 27 subjects who misarticulated the [r] phoneme and 38 subjects who misarticulated the [s] phoneme.

Source of Variation	*df*	*ms*	*F*	$F_{.05}$
Between Groups	1	26.90	.790	4.00
Within Groups	63	33.80		
Total	64			

The analysis of variance to evaluate the mean difference in pitch discrimination scores between the 27 subjects who misarticulated the [r] phoneme and 38 who misarticulated the [s] phoneme is summarized in Table 2. A two-part analysis was performed because the original analysis of variance test revealed considerable error variance, indicating that the matching variables were ineffective. Results of this analysis failed to show a significant difference between the two groups. The mean pitch discrimination score of the [r] group was 12.19 compared with a mean of 11.37 for the [s] group.

Conclusions and Discussion

The results of this study are in agreement with those of both Mange (*12*) and Bradley (*1*). This is of particular interest since a different pitch discrimination test was employed in each of the three studies. Of interest also is the fact that subjects in the present study were receiving speech therapy. Mange's subjects had not received speech therapy. Bradley did not report in his study whether or not his 11 subjects were receiving speech therapy. Speech therapy, with its emphasis on auditory discrimination training, might have been expected to have improved auditory perceptual skills.

The findings of the present study suggest the possibility that children who have functional articulation problems of different types may be, in general, poorer in pitch discrimination than normal children. This generalization is supported by the lack of significant differences between mean pitch discrimination scores for subjects who misarticulated [r] and those who misarticulated [s] and the significant difference between the entire misarticulation group and the control group.

The high residual variance, which can be seen in Table 1, indicated that IQ, sex, and grade were poor matching variables. The correlation between matched pairs of subjects was determined and found to be .135. The decision to use a matching design was made originally by the authors to utilize any advantage gained by controlling small portions of the total error variance. Early studies have indicated that IQ and pitch discrimination are correlated between .11 and .30 (*5, 13*). Although Seashore (*14*) maintained that no significant sex differences existed on factors such as pitch discrimination, and this opinion was supported by a study by Gilbert (*6*), an examination of Mange's original data in which he used sex as a matching variable indicated a favorable relationship between matched pairs on pitch discrimination scores. This appeared to be true as well for grade as a matching variable. For these reasons subjects in the present study were matched on IQ, sex, and grade. In essence the use of these matching variables functioned to decrease the efficiency of the experimental design, since the difficulties of matching subjects reduced the over-all number of subjects used and also reduced the degrees of freedom for the error term. It is interesting to note that studies of Eisenson, Kastein, and Schneiderman (*4*) and of Bradley (*1*) did not attempt to match subjects.

Summary

This study investigated pitch discrimination in school children with functional articulation errors in grades three through 12. A total of 65 subjects having articulation errors on either [r] or [s] were matched with a comparable group of normally speaking subjects on the basis of IQ, sex, and grade. The pitch subtest of the Tilson-Gretsch Music Test was administered to each subject following standard procedures.

Children with functional misarticulations were found to be poorer in mean pitch discrimination than normals. No evidence was found of a difference between the group misarticulating [r] and the group misarticulating [s] on

mean pitch discrimination scores. Matching variables of IQ, sex, and grade were found to be ineffective in increasing the precision of the experimental design.

References

1. Bradley, W. H., Some relationships between pitch discrimination and speech development. *Laryngoscope*, 69, 1959, 422-437.
2. Carrell, J., and Pendergast, Kathleen, An experimental study of the possible relation between errors of speech and spelling. *J. Speech Hearing Dis.*, 19, 1954, 327-335.
3. Carter, Eunice T., and Buck, M., Prognostic testing for functional articulation disorders among children in the first grade. *J. Speech Hearing Dis.*, 23, 1958, 124-133.
4. Eisenson, J., Kastein, Shulamith, and Schneiderman, Norma, An investigation into the ability of voice defectives to discriminate among differences in pitch and loudness. *J. Speech Hearing Dis.*, 23, 1958, 577-582.
5. Fracker, G. C., and Howard, V. M., Correlation between intelligence and musical talent among university students. *Psychol. Monogr.*, 39, 1928, 157-161.
6. Gilbert, G. M., Sex differences in musical 'aptitude' and training. *Psychol. Bull.*, 37, 1940, 508-509.
7. Gilkinson, H., The Seashore measures of musical talent and speech skill. *J. appl. Psychol.*, 27, 1943, 443-447.
8. Henderson, Florence M., Accuracy in testing the articulation of speech sounds. *J. educ. Res.*, 31, 1938, 348-356.
9. Hollingworth, Leta S., Musical sensitivity of children who test above 135 I.Q. (Stanford-Binet). *J. educ. Psychol.*, 17, 1926, 95-109.
10. Irwin, O. C., and Chen, Han Piao, A reliability study of speech sounds observed in the crying of newborn infants. *Child Develpm.*, 12, 1941, 351-368.
11. Kronvall, E. L., and Diehl, C. F., The relationship of auditory discrimination to articulatory defects of children with no known organic impairment. *J. Speech Hearing Dis.*, 19, 1954, 335-338.
12. Mange, C. V., Relationships between selected auditory perceptual factors and articulation ability. *J. Speech Hearing Res.*, 3, 1960, 67-74.
13. Ross, V. R., Relationships between intelligence, scholastic achievement, and musical talent. *J. juv. Res.*, 20, 1936, 47-64.
14. Seashore, C. E., Why no great women composers? *Mus. Educ. J.*, 26, 1940, 21, 88.
15. Templin, Mildred C., Certain language skills in children, their development and interrelationships. *Inst. Child Welf. Monogr.*, Ser. 26. Minneapolis: Univ. Minn. Press, 1957.
16. Tilson, L., and Gretsch, F., *The Tilson-Gretsch Musical Aptitude Test.* Chicago: Gretsch Mfg., 1941.
17. Travis, L. E., and Davis, Mildred G., The relation between faulty speech and lack of certain musical talent. *Psychol. Monogr.*, 36, 1927, 71-81.
18. Travis, L. E., and Rasmus, Bessie, The speech sound discrimination ability of cases with functional disorders of articulation. *Quart. J. Speech*, 17, 1931, 217-226.

Distance and Fingerspelling

HENRY M. MOSER

JOHN J. O'NEILL

HERBERT J. OYER

EDWARD A. ABERNATHY

BEN M. SHOWE, JR.

Dactylology or fingerspelling, a means of communication for the deaf, may have possibilities for use in signal situations such as those encountered on flight lines, in construction work, and on railways. Because these possibilities have not been fully explored, some preliminary testing of visual recognition of fingerspelled messages was necessary to determine the feasibility of this method of communication when distance is a factor. This report deals with four series of tests designed for this purpose: two indoors, in which the important variable was distance; two outdoors, in which the variables included not only distance but also differences in abilities and physical positions of both the senders and the receivers.

If the letters of the manual alphabet are observed from a variety of angles, it becomes obvious that from certain angles the distinguishing features of various letters show up more clearly. Presumably it is unlikely, in the normal usage of the manual alphabet, that all letters are transmitted at the angle which distinguishes them most effectively from other letters. This may be noted in comparing the fingerspelling of the alphabet when the hand is held close to the body, the so-called normal position (Figure 1), and when the hand is extended to the side (Figure 2). The normal positions have been described by Long (*1*) and in papers by other

Henry M. Moser (Ph.D., University of Iowa, 1936) is Director, Speech and Hearing Clinic, Ohio State University, and Project Supervisor, Air Force Contract No. 19(604)-4575, Psycholinguistics Laboratory, Ohio State University Research Foundation. John J. O'Neill (Ph.D., Ohio State University, 1951) is Professor of Speech, University of Illinois. Herbert J. Oyer (Ph.D., Ohio State University, 1955) is Associate Professor of Speech, Michigan State University. Edward A. Abernathy (Ph.D., Ohio State University, 1938) is Superintendent, Ohio School for the Deaf. Ben M. Showe, Jr. (M.A., Ohio State University, 1955) is Teacher-Librarian, Ohio School for the Deaf. Dr. O'Neill, Dr. Oyer, Dr. Abernathy, and Mr. Showe are former Research Associates, Psycholinguistics Laboratory, OSURF. This study was carried out under Air Force Contract No. 19(604)-1577, monitored by Operational Applications Laboratory, Air Research Development Command, Bolling Air Force Base, Washington.

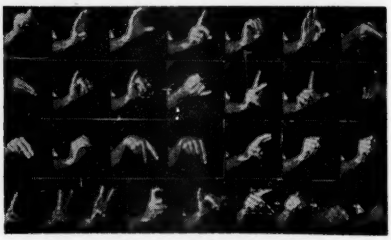

FIGURE 1. The fingerspelling alphabet presented in so-called normal manner with hand in front of and close to the body.

accepted authorities on the sign language. The literature, however, also mentions that the palm of the hand should face the person reading the message, a position which of course is difficult if the hand is held in front of and centered at the chest. It is not too unnatural, however, to spell with the forearm extending outward, perpendicular to the body, and in this way to achieve fairly easily the 'palm-to-the-reader' position.

The same letters transmitted by different persons, especially at some distance, may vary sufficiently to cause readers to confuse the stimulus letter with other letters. The tilt of the hand, excessive hand or wrist motion, incomplete closures, carelessness in delivery, even the shape and size of the hand may alter the appearance and clarity of many letters.

FIGURE 2. The fingerspelling alphabet presented with the arm extended to the right of the body.

Pilot Study

In order to discover weaknesses in the proposed experimental design and procedures of the present study, a pilot study was undertaken. The testing material for the study was composed of 26 nonmeaningful triplets of letters. Each letter of the alphabet was represented once in initial, medial, and final positions. Three packs of cards containing the alphabet letters were shuffled to get the letter combinations. This procedure was repeated until 16 different test lists were constructed. Six of these lists were used in the pilot study. Two male senders, both proficient fingerspellers, spelled the test triplets for the subjects.

The subjects were 25 readers selected from 40 students of grades seven through 12 at the Ohio School for the Deaf. Each prospective reader was given a trial test of 10 nonsense triplets. Those selected performed well on this test, wrote legibly, and had 20/20 vision.

The pilot study was conducted in the school gymnasium with the maximum illumination available from the overhead lights. Each of the senders presented the fingerspelling signals at three distances (50, 75, and 94 feet) with his hand against a background of black construction paper. A metronome set at 88 beats per minute enabled the senders to maintain a uniform rate of one letter per beat. Because the subjects were deaf and could respond in the main only to visual signals, it was difficult to control the time between the three-letter combinations.

Discussion. Several undesirable factors in the pilot study were noted. Seating arrangement of the subjects, in a single semicircle facing the sender, produced distortions in what they saw and variations in their responses. Room size limited to 94 feet the distance that signals could be sent. As the sender moved, there was variation in the amount of light reflected from his hand. The senders observed that since the readers were normally accustomed to seeing signals presented in front of the chest, the arm-to-the-side position was unnatural to read as well as to send.

Indoor Series 1

Procedure. For the actual test setting, a wide hallway was chosen with the sender seated on a low bookcase in a room opening off the end of the hallway. The sender did not alter his position and the illumination was kept constant by the use of two floodlights as well as the overhead lights. The two senders wore blue denim work aprons and presented the signals in the more natural hand-in-front-of-chest position.

There were 24 readers, all of whom had participated as subjects in the pilot study. They were seated in four groups of five and one group of four that were located 20, 50, 75, 100, and 125 feet, respectively, from the sender. After 'reading' two tests, one from each sender, as in the pilot study, each group of readers moved back one interval except the group at 125 feet which moved to the front position, 20 feet from the sender. The senders alternated in sending and in serving as a monitor for each other.

The 10 tests were given over a period of a week with four lists presented on each of the first two nights and two lists presented two nights later.

Results and Discussion. Although the letters were delivered in triplets the responses to them were scored individually and visual intelligibility values were computed on the basis of a total of 78 letters for each test. Confusion charts were constructed to indicate any trends evident in letter substitution.

Since the tests were run on three different nights, the first interpretation of the data involved a determination of the per cent recognition of the fingerspelling for each night of presentation. Thus it was possible to determine whether training or varying personal conditions not controlled by the experimental design had affected the test performance. There were only slight differences among the performances (92.05%, 92.08%, 93.2%) of the readers on the three test nights.

Average per cent intelligibility scores were computed for the test in terms of distance. At 20 feet the average in-

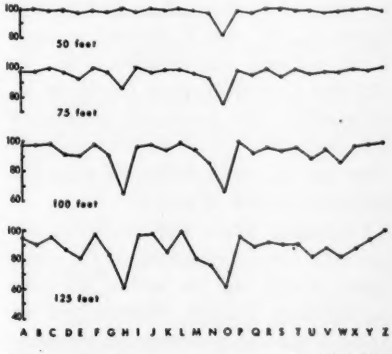

FIGURE 3. Average per cent intelligibility scores of 24 subjects for reading the fingerspelling alphabet at four distances indoors.

TABLE 1. Predominant substitutions by 24 readers for least intelligible alphabet letters at four distances in Indoor Series 1.

Stimulus Letters	*Substitutions*			
	at 50 feet	*at 75 feet*	*at 100 feet*	*at 125 feet*
O	E	E	E	E,S
H		G	G	G
E		S		S
W			V	V
N			M	M
U			R	R
M				N
G				H,Q
K				R

telligibility score was 99.2%; at 50 feet, 98.3%; at 75 feet, 96.2%; and at 125 feet, 87.6%. These scores indicate that there was a rather high intelligibility for the finger signals, especially at the shorter distances.

The intelligibility measures for the various letters at each of the four distances are represented in Figure 3. At 50 feet the intelligibility value of the letter O dropped sharply. At 75 feet this decrease in intelligibility was again evident for O along with a decrease in the intelligibility for the letters H and E. At 100 feet the per cent intelligibility value of O dropped to 65.3%. The letters N, W, and U showed a somewhat reduced intelligibility also. At 125 feet O and H were lowest in intelligibility. The previously mentioned less intelligible letters at 100 feet again were low in intelligibility, as were the additional letters M and G.

Confusion tables indicate predominance of certain substitutions for those letters which had the lowest per cent intelligibility ratings. Summarized in Table 1 are the predominant substitutions for the stimulus letters of lowest intelligibility. For O, the letter most

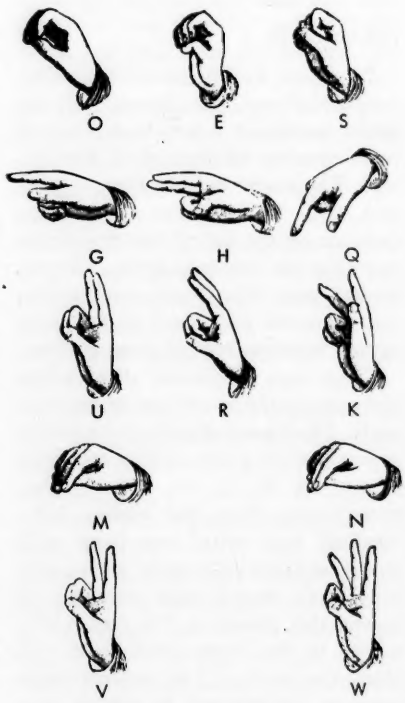

FIGURE 4. Groups of hand signals most often confused by 24 subjects in reading fingerspelling at various distances indoors and outdoors. Each line is one confusion group.

TABLE 2. Summary of analysis of variance for evaluating differences among distances and between senders for Indoor Series 1.

Source	*df*	*ms*	*F*
Distances (D)	4	1289.05	
Senders (S)	1	91.26	
Observers (O)	23	34.35	
DS	4	16.67	2.65*
DO	92	11.41	
SO	23	4.70	
DSO	92	6.29	
Total	239		

*$F_{.05} = 2.45$.

frequently confused, the substitute response was E almost consistently, although S also was substituted frequently at 125 feet. At 100 and 125 feet, V, M, and R were most frequently substituted for the stimulus letters, W, N, and U. At 100 and 125 feet, R was the predominant substitution for U, and at 125 feet was also the predominant substitution for K. Two letters were highly confused with several letters in the alphabet but not enough with any one letter to be included in Table 1. The letter P was frequently used as a substitute for both H and Q while the letter R was used as a substitute not only for U and K but also for V and D. The highly confused hand signals are shown in Figure 4. The letters J and Z had the greatest intelligibility. This might be expected in that they are the only letters which involve motion and there are distinct differences between them.

Analysis of Variance. Results of a triple analysis of variance using the factors (a) distance, (b) senders, and (c) observers, are presented in Table 2. The criterion measure was the total number of correct responses to one sender for one observer at one distance. Because of the significant interaction between distances and senders, a sepa-

TABLE 3. Summary of analysis of simple effects of distances and senders in Indoor Series 1.

Source	*df*	*ms*	*F*	*p**
Distance 1 (20 feet)	1	.33	.07	ns
Distance 2 (50 feet)	1	2.52	.54	ns
Distance 3 (75 feet)	1	21.33	4.54	5%
Distance 4 (100 feet)	1	102.08	21.71	1%
Distance 5 (125 feet)	1	31.69	6.74	5%
Sender 1	4	543.82	47.66	1%
Sender 2	4	761.91	66.77	1%

*Point in *F* distribution.

rate test was made for each of the five distances and each of the two senders. The results of these tests are presented in Table 3. The differences among distances for each of the senders were significant at the 1% level. Differences were in favor of Sender 1 at all but the first distance, 20 feet, and at the fourth distance, 100 feet, this difference was significant. All fingerspelling became significantly less intelligible at each increase of distance after the second distance, 50 feet.

Indoor Series 2

The high intelligibility scores at 125 feet indicated that even greater distances should be tested. Because of time and space curtailments, the five best readers (Group I) selected according to their performances at 125 feet, were tested with the pilot study test lists at 175, 200, and 250 feet. The five next best readers (Group II), selected in a similar manner, were tested at 150, 175, and 225 feet. In comparison with Group II, Group I had a less erratic pattern response from distance to distance and also performed better at 175 feet. The result of a t test for evaluation of the mean difference between the two groups at this distance was significant at approximately the 6% level ($t = 2.66$; $t_{.05}$ with 4 $df = 2.78$). Group I, reading at 175, 200, and 250 feet, performed better at each of these distances than Group II reading at 150, 175, and 225 feet, respectively. Even at 250 feet, Group I had an average intelligibility score above 50%. In this latter testing at extreme distances the intervening overhead hall lights were turned off to help direct the attention of the group toward the signal source. In the previous testing the readers had reported that at the greater distances these lights were disturbing.

Outdoor Series 1

In order to evaluate the reliability of the results of the indoor series, and also to test the effect of even greater distances on letter intelligibility, two series of tests were conducted outdoors. They were performed on different days and had only two test distances in common.

The method of constructing the tests was similar to that employed indoors and the same two senders served in both situations. The readers were the 10 best observers used in the indoor testing for the longer distances, that is, the same as Groups I and II above. The four girls in these groups had returned from camp just prior to the first outdoor testing and complained of being tired. There were no fatigue factors of an unusual sort involved in the second series of tests.

Procedure. The tests of Outdoor Series 1 were run in early June, between the hours of 4 and 6 p.m. The day was bright with scattered cirrus clouds. An illumination meter, Weston Model 603, was used to obtain readings of light reflected off the hand at 10 inches. The readings for these tests were 200 to 300 foot-candles. The senders wore dark blue denim bib-aprons over their white shirts to provide a good background for the spelling hand while transmitting test words from a natural position in front of the chest. The readers were provided with lapboards, answer sheets, and pencils. They were seated in the shade in two rows with five students in each row.

TABLE 4. Most frequent substitutions by 10 superior readers for the least intelligible alphabet letters at various distances in Outdoor Series 1.

Stimulus Letters	*Substitutions*				
	at 125 ft	*at 150 ft*	*at 175 ft*	*at 250 ft*	*at 300 ft*
E	O,S	S	O,S		O
H	G	G	G	G	G,Q
N	P	Q,P	P,Q	Q,M	P,Q
O	E	E	E,S		
Q	P	P	P		
W	V	V	V		V
G		Q			P
M		N	N	N	N,P,Q
R		C	X	C,D	
U		R	R	R	R
T			S,A		A
F				I	I
K				D	R
S					
V				O,E	
B				W	R,U
Y					A

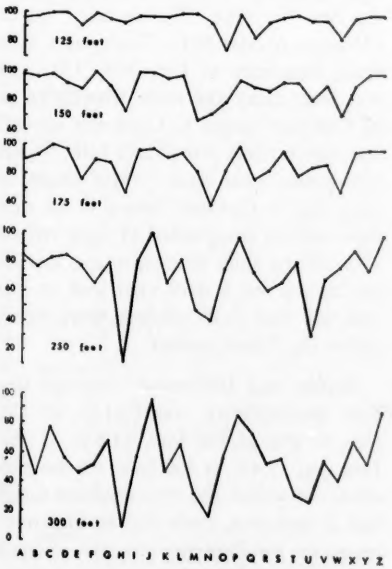

FIGURE 5. Average per cent intelligibility scores for 10 superior readers in reading the fingerspelling alphabet outdoors (Series 1).

The senders stood in the sunlight facing the sun. Test signals were examined at five distances: 125, 150, 175, 250, and 300 feet. The sender alerted the subjects for each set of signals by raising his hand over his head.

Results and Discussion. At 125 feet the average intelligibility was 95.9%; at 150 feet, 90.3%; at 175 feet, 88.1%; at 250 feet, 72.9%; and at 300 feet, 56.3%. These results appear to be good but must be interpreted with the consideration that previously determined superior readers served as subjects.

Average per cent intelligibility values for each alphabet letter are shown for each test distance in Figure 5. The most frequently used substitutions for those alphabet letters with low intelligibility are presented in Table 4. All alphabet letters which had an intel-

ligibility value below the average intelligibility value of the entire alphabet at that distance are presented in this table. The substitution patterns which appear consistently throughout all the test distances are O and S for E; G for H; P and Q for N; E, and occasionally S, for O; P for Q; V for W; N, and occasionally P and Q, for M; and R for U. In some cases the confusions are reciprocating (for example, P, Q, and N as substitutions for each other in multiple relationship) and in others they are not (for example, G as a substitution for H but not the converse).

These graphic representations do not necessarily present a clear picture of the confusions involved. Examination of confusion tables showed that for some stimulus letters heavy concentrations of confusion occurred which were not sufficient to cause intelligibility scores low enough for them to be included in Table 1. For some stimulus letters, therefore, additional interpretation from the confusion tables seems important. The O and S substitution for E at 125 feet is not as concentrated as the V for W, the E for O, the P for Q, or the G for H. At 150 feet the G appears to be confused with many signs other than H, and P and S are not only frequent substitutions at 150 feet but also are used frequently at 175 feet. These same letters, along with Q and R, are used frequently as substitutes for many letters at 175 feet. As the distance increases, the scatter of substitution responses increases. At 250 feet, R is a major substitution response, occurring with a frequency of five or more for B, D, K, and U; Q occurs as a substitution response for G, H, M, and N; A is a frequent substitution as a response to the signals I, T, and Y. At 300 feet, Q and R are again favorite substitutions; R occurs frequently as a substitute for B, D, F, K, and U; Q occurs most often as a substitute for G, H, M, N, P, and X; A occurs for I, S, T, V, and Y; P occurs for G, H, M, N, and Q.

In this outdoor experiment at various distances, G, P, S, Q, R, and A are most frequently substituted for other letters of the fingerspelling alphabet.

Outdoor Series 2

Procedure. The testing occurred at the same time of day as for Outdoor Series 1. The clouds were more scattered on this day and the reflected light was measured as 300 to 400 footcandles by the illumination meter (Weston Model 603). Tests were sent from distances at 175, 300, 350, and 400 feet, using the same procedure as in Outdoor Series 1. Until the second test, the senders wore dark blue denim bib-aprons over their white shirts as they did in Outdoor Series 1. At that time readers complained of light reflection off the shirt sleeves, so the second sender put on a dark coat and at 350 and 400 feet both senders wore coats under the denim aprons.

Results and Discussion. Average per cent intelligibility was 93.1% at 175 feet, 63.9% at 300 feet, 53.8% at 350 feet, and 37.4% at 400 feet. At the two distances which the two outdoor series had in common, there was an improvement in intelligibility in the second series: at 175 feet, the average per cent intelligibility scores for the second series were 5% higher than those for

TABLE 5. Most frequent substitutions by 10 superior readers for least intelligible alphabet letters at various distances in Outdoor Series 2.

Stimulus Letters	*Substitutions*			
	at 175 feet	*at 300 feet*	*at 350 feet*	*at 400 feet*
E	S	O	X	S,X
H	G	P,G	G	G
N	Q,P	Q,P	Q	Q
R	X	C		D
W	V	V		B,K
D		X,R	R	R
I		O,E	E	S
K		R,D	R,D	R
M		N	N	Q,N
O		E	S	E,S
S		A		
T		A	A,S	A
U		R	R	R
V		U		R,X
X			R	B,R
L				C
Q			P	

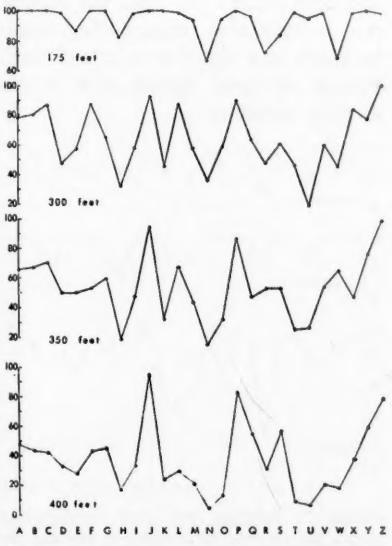

FIGURE 6. Average per cent intelligibility scores for 10 superior readers in reading the fingerspelling alphabet outdoors (Series 2).

the first (93.1% compared with 88.1%) and at 300 feet the scores for the second series were 7.6% higher than those for the first (63.9% compared with 56.3%). The fatigue mentioned by the girl readers in the first series may have caused this difference in results. The fact that the senders wore dark coats to eliminate glare off their shirt sleeves might also be a factor. Figure 6 presents the intelligibility values for the individual alphabet letters at each of the test distances.

In Table 5 are presented the most frequent substitutions for those letters with per cent intelligibility scores below the average. Recurring substitutions were G for H; Q and P for N; R for D; R for K; N for M; S and E for O; A for T; and R for U. The confusion tables indicate that at 300 feet R is a common substitution for

U, K, D, and B; P is a substitution for G, H, N, and Q; K is a frequent substitution for U, V, and W. At 350 feet, R again heads the list as a favorite substitution, occurring for X, V, U, K, D, C, and B; X is used as a frequent substitution for E, G, H, O, and R; S is a substitution for A, E, O, and T. At 400 feet, the letters most used as substitutions are R, Q, P, X, B, and S: R occurs as a substitution for X, U, T, N, M, H, and G; Q occurs for N, M, H, and G; P for Q, N, M, and H; X for V, T, E, and C; B for X, W, and F; and S for T, O, I, E, and A. Summarizing, it is found that R, P, K, X, and S are frequently used as substitutions throughout the alphabet.

It was considered unwise to attach any great importance to the data concerning the distances at which the average intelligibility was below or only slightly above 50%. There would naturally be a wider range of guesses involved here than at the shorter distances.

Conclusions

A comparison of the average per cent intelligibility scores in the indoors and outdoors performances of the same subjects, that is, of the group reading in the second indoors series and in both outdoors series, indicates that intelligibility is higher in the outdoors performances than in the indoors series and that in the second outdoors it is superior to the first (Figure 7). The mean scores were 79.6% for Indoor Series 2; 88.1% for Outdoor Series 1; and 93.1% for Outdoor Series 2.

It is possible that these differences in scores reflected not only the ability of the senders but also the readers' familiarity with the senders' characteristic manners of spelling. It is possible also that the senders, knowing what letters were causing confusion, made a conscious effort to send these particular letters with greater clarity. In addition to distance, the influence of light may have been a factor inasmuch as the light difference was one of the great differences between the indoor and the outdoor series.

A certain consistency was noted in both the letters which were most often misread and in the substitutions most often read for them. It was noted further that substitutions tended to follow the gross hand shape; for instance, closed hand signals were confused with closed hand signals, open hand signals with open hand signals.

Since, under the conditions of the experiment, the intelligibility of fingerspelling was found quite high at distances around 175 feet, the suggestion is made that fingerspelling might be usable as a supplement to and clarification of hand signals now in use at such distances.

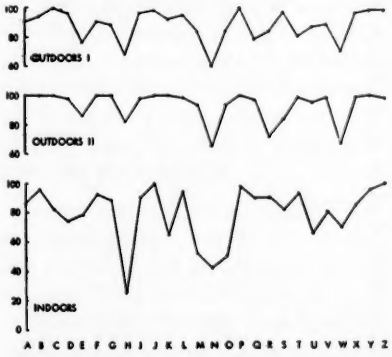

FIGURE 7. Average per cent intelligibility scores for 10 superior readers in reading the fingerspelling alphabet presented by the same two senders at 175 feet, once indoors and twice outdoors.

Summary

A test was made of the visual intelligibility of nonmeaningful alphabet triplets presented in fingerspelling to 24 deaf students under conditions of artificial lighting and natural lighting. Two senders, familiar with the fingerspelling alphabet and known by the subjects, presented the materials.

Results indicated that distance between the sender and the receivers was the prime factor affecting intelligibility. At 125 feet, the greatest distance tested with the entire group, the average intelligibility was 87.6%. At 250 feet, the five best readers were able to recognize better than 50% of the materials. Outdoors the average per cent intelligibility was higher than for the corresponding distance indoors and the average intelligibility was above 50% at 300 feet. Under all conditions of testing there were tendencies for certain letters to receive certain substitutions and other letters to be highly confused with a wide variety of letters. Substitutions tended to follow gross hand shape.

In general, the intelligibility of alphabet letters was sufficiently high up to 175 feet to indicate possibilities of using the fingerspelling alphabet as a means of supplementing or clarifying present hand signals now in use.

Reference

1. Long, J. S., *The Sign Language*. (2nd ed.) Des Moines: Henderson, 1918.

► RESEARCH NEWS NOTE

White Noise and Stuttering

► A number of investigators have observed reduction of stuttering when the acoustic environment of the speaker is altered. Among such alterations is the use of high intensity white noise which presumably 'masks' or prevents the speaker from hearing the sounds of his own speech. In a recent study, Cherry and Sayers (*J. psychosom. Res.*, 1956, 233-246) demonstrated a total inhibition of stuttering during exposure to high intensity white noise. When the white noise was filtered to pass or reject frequencies below 500 cps, the inhibition was more effective with the low pass white noise; that is, when the noise contained a predominance of low frequencies which are in the region of the spectrum at which most of the energy of speech is concentrated. This leads to the implication that noise achieves its effectiveness for the stutterer by reducing his hearing of the sounds of his own speech.

In the course of an extensive series of experiments in our laboratory on the role of a variety of acoustic conditions—white noise, pure tones, delayed auditory feedback—in the reduction of stuttering, we came to doubt that noise achieves its effect by reducing the ability of the stutterer to hear the sounds of his own speech.

We undertook the following means of studying this problem. The presentation of white noise during oral reading was triggered by a voice-actuated relay. In one condition of the experiment, presentation of the noise occurred only during the phonation of speech sounds. During silent periods, no noise was presented. In a second condition, the noise was presented only during the silent periods, and not during the phonation of sounds. In this condition, therefore, the noise could not mask the speaker's hearing of his own speech sounds. These two conditions were compared with (a) continuous noise, as is normally used in 'masking' experiments; (b) wearing of earphones but no noise presented; and (c) several control readings under normal conditions. If the reduction of stuttering is due to the fact that the noise prevents the stutterer from hearing his own speech, the condition in which noise is present only during phonation should be as effective as continuous noise, while the condition in which noise is present only during silent periods should be relatively ineffective.

In the following table are the results in terms of correct words per second for nine stutterers for the five conditions: average control (AC), phones only (PO), noise during silence (NS), noise during phonation (NP), and continuous noise (CN). Each row presents the scores for one subject.

AC	*PO*	*NS*	*NP*	*CN*
.33	.37	.28	.24	.86
.41	.37	1.95	2.14	2.37
.60	1.55	2.58	1.70	1.96
.61	.77	2.05	1.94	1.83
1.12	1.59	1.87	1.51	1.28
1.45	1.07	1.90	2.21	2.12
1.68	2.15	1.81	1.76	2.22
1.70	1.53	2.72	1.79	2.29
1.87	2.42	1.92	2.35	2.09
Mean				
1.09	1.31	1.90	1.74	1.89

$\chi_r^2 = 12.53$; $p < .02$.

The noise conditions clearly improve the correct word rate as may be seen in the rise from mean rates of 1.09 and 1.31 under the no-noise conditions to mean rates of 1.90, 1.74, and 1.89 under the noise conditions ($p < .02$, Friedman two-way analysis of variance). However, when comparing the three noise conditions, it may be seen that the condition in which noise occurs only during silent periods, in other words the conditions in which noise cannot mask the speaker's hearing of his own speech, is just as effective as the other two conditions in which masking does occur.

It seems clear that a re-evaluation of the way in which noise reduces stuttering is in order.

Samuel Sutton, Ph.D.
Columbia University
Richard Allen Chase, M.D.
Yale University School of Medicine

Low Sensation Level Delayed Clicks and Keytapping

RICHARD ALLEN CHASE

SAMUEL SUTTON

EDMUND P. FOWLER, JR.

THOMAS H. FAY, JR.

HOWARD B. RUHM

Hearing is normally tested by presenting pure tones to a subject, who reports his hearing experience. Although this technique gives accurate clinical information, its dependence on report of subjective experience limits its validity when applied to several groups of individuals who are either unable or unwilling to give such a response. One such group are the children brought for medical evaluation because they do not respond appropriately to verbal communication and for whom the differential diagnosis often includes middle ear and sensorineural hearing loss, childhood schizophrenia, and organic brain damage. Another type of person who cannot be tested by routine techniques is the malingerer.

A technique introduced recently for gaining additional information with respect to the hearing acuity of these problem groups is delayed auditory feedback. Auditory stimuli produced by motor activity and used to monitor motor activity are called auditory feedback. An example of auditory feedback is the sound of a person's own voice as he speaks. A delay in the return of a speaker's voice to his ears, delayed auditory feedback, results in marked changes in his speech. He speaks louder, slows his rate of speaking, and repeats sounds (*4*). Many attempts have been made to estimate the degree of hearing impairment by noting the lowest in-

Richard Allen Chase (M.D., Columbia University College of Physicians and Surgeons, 1960) is Research Fellow, Department of Otolaryngology and Division of Medical Psychology, Department of Psychiatry, Columbia University; Samuel Sutton (Ph.D., University of Chicago, 1955) is Associate Research Scientist, Biometrics Research, New York State Department of Mental Hygiene, and Research Associate, Department of Psychiatry, Columbia University; Edmund P. Fowler, Jr. (M.D., Columbia University College of Physicians and Surgeons, 1930) is Professor of Otolaryngology, Columbia University College of Physicians and Surgeons; Thomas H. Fay, Jr. (Ph.D., University of Illinois, 1958) is Supervisor of Audiology, Columbia-Presbyterian Medical Center; Howard B. Ruhm (Ph.D., Northwestern University, 1959), formerly Research Audiologist, Veterans Administration Audiology Research Program, is Director of Audiological Research, University of Oklahoma. This work was supported under grants B-1686, B-2699, and M-1541, from the National Institutes of Health.

tensity of the delayed auditory feedback which will produce changes in speech (*1*, *2*, *5*, *8*, *9*, 7). The fact that speech consists of many frequencies with sound energy distributed unevenly throughout these frequencies has made it impossible to obtain precise information about hearing acuity by this means. Another factor which decreases the usefulness of the delayed auditory feedback of speech for auditory testing is the difficulty of obtaining a sensitive and unambiguous measure of the complex alterations which occur in speech under delayed auditory feedback. In one study (*8*), a panel of judges was used to decide whether a subject was speaking under delayed or normal feedback. Majority agreement was correct in only two out of 10 cases when the delayed feedback was at 10 db sensation level. The use of reading time as a measure gave only one correct identification out of a possible 10 at this sensation level.

The authors have recently shown that a delayed click presented to a subject's earphones each time he taps on a key results in changes in keytapping which are analogous to changes in speech under delayed auditory feedback (*3*). The usefulness of this technique in testing hearing depends on whether unambiguous changes in tapping performance can be obtained at an intensity level which is close to the threshold of hearing. It is the purpose of this report to present data pertinent to this question.

Procedure

The subjects were 20 young adults with normal hearing. Each subject's

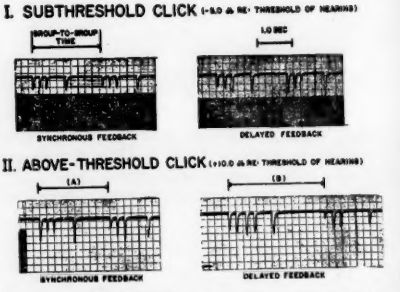

FIGURE 1. Graphic time-amplitude characteristics of groups of four taps on a key: I, subthreshold click (no change in group-to-group time with delay); II, above-threshold click (increase in group-to-group time with delay).

threshold of hearing was determined for a click of 400 μsec duration and containing frequencies from 500 to 2000 cps. The subject was then asked to tap on a key according to the following pattern: three taps, pause, one tap (. . . .). The subject was to continue to tap such groups of four, with a somewhat longer pause between groups, until told to stop. Each subject picked his own rate, but was told to be as constant as possible in the rate and pressure of his tapping. Each time the subject hit the key, the output of a Baldwin Lima Hamilton, Model A-7, SR-4 strain gauge was used to give a continuous record of the time and amplitude characteristics of his tapping (Figure 1). A schematic diagram of the equip-

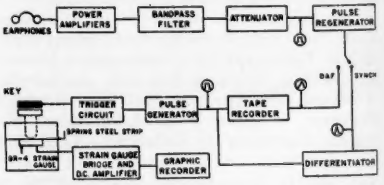

FIGURE 2. Schematic diagram of equipment.

ment is shown in Figure 2. Each time the subject tapped on the key, he triggered presentation of the click (described above) to his earphones. Special circuits were designed to insure that a square-wave pulse of 400 μsec was presented at the attenuators whether the click was delayed by use of a tape recorder or whether the click was returned without delay to the subject's earphones.[1] The general arrangement of the circuits used to provide delayed and undelayed clicks is shown in Figure 2. The clicks could be attenuated in 1 db steps from zero to 100 db below the threshold of normal hearing, using a Hewlett-Packard #350 B attenuator. Two cascaded United Transformer filters (types HML and IML) were used to obtain a pass band of 500 to 2000 cps for the click delivered to the earphones. The power amplifier was a 60-watt Acrosound Mark II, and the earphones were Permoflux PDR-1.

In the first condition of the experiment, the clicks were presented at 5 db below the subject's threshold of hearing. The subject was instructed to continue tapping in groups of four according to the prescribed pattern until told to stop. Without pause or warning, the experimenter alternated from synchronous feedback of the clicks to delayed feedback of the clicks after every 12th group of taps to obtain three samples of 12 groups each, two samples under synchronous feedback and one under delayed feedback. In the second condition of the experiment, the same procedure was followed, but the intensity of the clicks was now 10 db above the subject's threshold of hearing and six samples of 12 groups each were obtained for comparison of synchronous and delayed feedback.

[1]The technique for obtaining delayed auditory feedback has been described elsewhere (*6*).

TABLE 1. Group-to-group measures, in millimeters of tape directly proportioned to time, for one individual (Subject 13) for the alternating synchronous (S) and delay (D) sequence of the above-threshold condition.

	S	*D*	*S*	*D*	*S*	*D*
Group						
4	45	55	48	52	51	53
5	45	54	50	57	52	50
6	50	55	50	57	52	51
7	48	55	49	54	55	51
8	49	55	50	53	48	50
9	49	66	51	58	48	53
10	49	55	51	58	50	52
Mean	47.9	56.4	49.9	55.6	50.9	51.4

The time in seconds from the first tap of one group to the first tap of the next group was defined as the group-to-group time (Figure 1). Means of these measures for groups 4 through 10 constituted the raw scores for the statistical analyses. Measures for the first three groups of each sample were omitted in the analysis of the data because the first few groups after an alternation tend to show a lag effect from the previous condition. One subject's record for the above-threshold condition is shown in Table 1. (The entries in the table are in millimeters of tape which is directly proportioned to time.)

Results

A Friedman two-way analysis of variance by ranks (*10*, p. 166) was performed, employing the ranks, for each of the 20 subjects, of the mean scores under synchronous and under delayed feedback trials for the subthreshold

condition. There were no significant differences among the two synchronous and one delayed feedback trials ($\chi_r^2 = 3.48$ with 2 *df*). The same analysis for testing differences among the six above-threshold sample means indicated highly significant differences ($\chi_r^2 = 53$ with 5 *df;* significant beyond the .001 level). The mean ranks for the three delayed feedback samples were considerably and consistently larger than the mean ranks for the three synchronous feedback trials. The evidence is thus strong that group-to-group time is longer under delayed feedback than under synchronous feedback.

Since the interest here was in diagnostic applications, a significance test was performed on each individual's data for the 10 db above-threshold feedback of the clicks. Because of the distribution and the narrow range of scores (Table 1), the Kolmogorov-Smirnov test (*10*, p. 127) was used to evaluate differences between synchronous and delayed feedback trials. For 15 of the 20 subjects the result was significant beyond the .05 level (χ^2 range: 7.55 to 42.00); for four of the remaining subjects, *p* levels ranged from .06 to .10; and for only one subject did the *p* level exceed .10.

Part of one subject's graphic record is shown in Figure 1. When the above-threshold click is presented, the group-to-group time increases under delayed feedback. This subject made an extra tap in his first group under delayed feedback with the above-threshold click. Such errors are commonly seen in keytapping records under delay, but the group-to-group time increases for all subjects in general even in the absence of errors in number of taps.

Discussion

It seems clear that the disturbance in motor performance which follows a delay in auditory feedback has potential for testing hearing acuity in subjects who are either unable or unwilling to give consistent responses to standard audiometric tests. However, a number of further studies are required before delayed feedback of the auditory accompaniments of tapping may be considered suitable for clinical use. First, pure-tone bursts need to be substituted for the click in order to obtain an audiometric curve. Second, the study should be repeated with populations having different etiological types of hearing loss. Harford and Jerger (*9*), using delayed speech feedback to study hard-of-hearing individuals with an etiology of endolymphatic hydrops, found that such individuals were more disturbed in their speech performance at medium and high intensity delayed auditory feedback than were normals. This is presumably due to the presence of recruitment in this disorder. Paradoxically, however, patients with otosclerosis who do not show recruitment also were more disturbed in their speech performance by high intensity delayed auditory feedback than were normals. Such factors suggest that the extrapolation of results with normal populations to hard-of-hearing populations, or the attempt to infer threshold level from above-threshold data, is beset with difficulties. In this connection, study of the delayed feedback of keytapping may help to elucidate some of the seemingly contradictory results obtained with delayed speech feedback, for example, the paradoxical effects obtained in patients with otosclerosis.

It may be that the results obtained with this group with delayed speech feedback are due to a complex relationship between air-conducted and bone-conducted feedback of the speech. Since the keytapping task produces no bone-conducted feedback, this may be used to test such possibilities.

A third consideration on which further study is needed is the nature of the keytapping task to be used. It will be noted that the delayed feedback of the task in the present study (... .) showed differences between synchronous and delayed feedback significant at the .05 level for only 15 of the 20 subjects. While this is a substantial achievement at 10 db sensation level, modification of the required motor task may improve the sensitivity of the technique. Another modification which might increase sensitivity is the use of more than one delay time of the auditory feedback. A great deal of research information is available which indicates that the delay time which causes most disturbance in speech differs from individual to individual. It is to be expected that similar interindividual differences for the keytapping task will be found.

Summary

A technique is suggested which may be useful along with present clinical methods of evaluating threshold of hearing. It consists of observing changes in the temporal characteristics of simple keytapping performance when a delay is introduced in the hearing of the click which accompanies each tap.

The 20 subjects tapped on a key in a simple temporal pattern under synchronous and delayed feedback of the click. The feedback was below hearing threshold in the first condition of the experiment, and at 10 db sensation level in the second condition. Differences between tapping performance under synchronous and delayed feedback in the subthreshold condition were not significant, but at the 10 db sensation level they were highly significant ($p < .001$). Intraindividual analysis of the above-threshold data yielded significant differences ($p < .05$) between the synchronous and delayed condition for 15 of the 20 individuals. A number of suggestions are made for adapting this technique to clinical use.

Acknowledgment

The authors gratefully acknowledge the assistance of Mr. Kai Klemm and Mr. George Guilfoyle.

References

1. Azzi, A., Un nuovo metodo per svelare la simulazione della sordità (A new method for revealing the simulation of deafness). *Arch. ital. Otol.*, 62, 1951, 329-334.
2. Butler, R. A., and Galloway, F. T., Performances of normal-hearing and hard-of-hearing persons on the delayed feedback task. *J. Speech Hearing Res.*, 2, 1959, 84-90.
3. Chase, R. A., Harvey, S., Standfast, Susan, Rapin, Isabelle, and Sutton, S., Comparison of the effects of delayed auditory feedback on speech and keytapping. *Science*, 129, 1959, 903-904.
4. Chase, R. A., Sutton, S., and First, Daphne, Bibliography: delayed auditory feedback. *J. Speech Hearing Res.*, 2, 1959, 193-200.
5. Elliott, Charlotte M., Responses of hard of hearing subjects to delayed side-tone. M.A. thesis, Vanderbilt Univ., 1955.
6. Fairbanks, G., and Jaeger, R., A device for continuously variable time delay of

headset. monitoring during magnetic recording of speech. *J. Speech Hearing Dis.*, 16, 1951, 162-164.

7. Gibbons, E. W., and Winchester, R. A., A delayed sidetone test for detecting uniaural functional deafness. *A.M.A. Arch. Otolaryng.*, 66, 1957, 70-78.
8. Hanley, C. N., and Tiffany, W. R., An investigation into the use of electromechanically delayed sidetone in auditory testing. *J. Speech Hearing Dis.*, 19, 1954, 367-374.
9. Harford, E. R. and Jerger, J. F., Effect of loudness recruitment on delayed speech feedback. *J. Speech Hearing Res.*, 2, 1959, 361-368.
10. Siegel, S., *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill, 1956.

► RESEARCH NEWS NOTE

Social Variables and Speech Disorders

► An extensive research project (estimated date of completion: 1963) will attempt to delineate the role of 'social variables' in the onset and development of speech disorders. The writer has received a grant from Michigan State University to initiate this project.

The research is concerned primarily with the epidemiology of speech disorders: the relative incidence and prevalence of diagnosed speech problems over the social class structure. That is, a study of speech disorders will be conducted through comparison of frequencies, patterns, and distributions of such disorders in socially stratified populations. The study will include also (a) an analysis of the relationships of specific diagnoses of speech disorders to the class structure (are organically based speech disorders more prevalent in the 'lower' segments of the social structure?); (b) a study of the role of 'social mobility' in the incidence of speech disorders; (c) an investigation of clinical treatment in relation to the class membership of the patient; (d) an analysis of the relations between various indices of social stratification (occupation, value system, etc.) and speech disorders.

A theoretical model has been derived from the extant literature pertaining to sociology, anthropology, psychology, and speech pathology. An interview schedule is being designed to study a group of parents of speech defective children and a control group of parents of normal-speaking children. The author will utilize Hollingshead's index of social class to evaluate the experimental and control groups. Relations between social class and speech disorders will then be examined by nonparametric statistical methods.

Lonnie L. Emerick, M.A.
Assistant Instructor, Speech Department
Michigan State University
East Lansing, Michigan

Maternal and Environmental Factors Related to Success in Speech Improvement Training

PHYLLIS BURGESS ANDERSLAND

The effectiveness of speech training for children depends upon its suitability for the individual in terms of the original and sustaining causes of the problem as well as upon the child's capacity for change. Lack of knowledge concerning etiological conditions and the difficulty of research which takes into consideration the variety of factors which may be involved have hampered clinical assessment and treatment. One experimental approach entails a comparison of background factors in children who improve after receiving speech training with those who do not. Differences between the groups may provide indications of those factors most relevant to success in speech training.

The purpose of this study is twofold: (a) to ascertain what relationships exist between children's articulation and certain factors thought (*10*, pp. 707- 768) to be related to the etiology of functional articulatory problems and (b) to determine in what ways kindergarten speech improvement training alters these relationships. A postdictive approach was utilized; that is, groups of children with known articulation development, some of whom had received speech training, were compared on a number of environmental factors. More specifically, an attempt was made to determine the relationship of certain maternal personality traits, maternal attitudes, and family socioeconomic level to improvement in articulation.

Wood (*14*) reported that functional articulatory errors in children and the effectiveness of speech therapy were significantly associated with maternal maladjustment. Certain methodological procedures in his study should be noted in an evaluation of these conclusions: First, the judgment of maternal maladjustment was based upon a comparison of his subjects with the normative population of the Bernreuter Personality Inventory and the California Test of Personality (CTP) rather than upon comparison with a control group selected for the immediate purpose. Second, lower cutting scores of 'neuroticism' than were validated were used for the tests, thus probably giving a biased picture of maladjustment.[1] Finally, while children of mothers receiving counselling manifested more

Phyllis Burgess Andersland (M.S., Purdue University, 1959) is psychologist for the East Lansing, Mich., public schools. This article is based on a master's thesis completed under the direction of Prof. John M. Hadley.

[1]Bernreuter (*3*, p. 1) reported that people scoring above the 98th percentile would probably profit from psychiatric or medical advice. Wood (*14*) considered that parents above the 50th percentile manifested neuroticism on the test. No validated cutting scores were supplied by the CTP (*12*, p. 6).

improvement in speech therapy than children of mothers who were not counselled, the interpretation of this finding as indicating the importance of maternal adjustment for successful speech therapy might be open to question. The mothers were given the opportunity to discuss their problems and the results from personality tests, but the counselling procedure as reported by Wood centered on the teaching of techniques for handling children. Nearly half of the negative home environment factors listed pertained to problems of child management. Since no change in maternal adjustment was demonstrated directly, accelerated speech improvement may have resulted from better parent-child relationships. These considerations suggest that further investigation and a general re-examination of this problem are indicated.

During the 1951-52 academic year Wilson (*13*) studied 242 children who were attending kindergarten in five public schools in Lafayette, Indiana. For a period of 12 weeks in their regular half-day school sessions, children in the experimental group received daily speech lessons from their regular kindergarten teachers. Children in the control group were taught by the same five kindergarten teachers during the other half-day sessions in the same period but received no speech lessons. The two groups were relatively homogeneous in terms of age, sex, intelligence, social maturity, and socioeconomic status. Results of the study indicated that the children who had received classroom speech improvement lessons showed significantly greater improvement in articulation than did the control group. The usefulness of kindergarten speech improvement lessons of the Wilson type was considered substantiated.

Procedure

Experimental Subjects. The effects of speech training in relation to 15 maternal and environmental factors were investigated through comparisons among the seven groups of mothers, the groups established solely on the basis of the speech histories of their children. Participating in the study were 94 mothers of children studied by Wilson (*13*). These mothers represented 85% of the possible usable subjects. (The remaining 15% included mothers from six of the seven groups and from all socioeconomic brackets.) These 94 mothers had remained in Lafayette, were willing to cooperate, and had children in one of the seven criterion groups.

Establishing Criterion Groups. The articulation of those children from Wilson's study who were available was reassessed by Drexler (*4*) and others during the 1957-1958 school year. Satisfactory reliability of articulation assessment was established by these studies (*13, 4*). Each child originally accepted had hearing and vision within the normal range and was without gross organic defect. The present investigator classified the children into groups according to whether or not they had received kindergarten speech improvement lessons and/or subsequent speech therapy. They were further classified according to the absence or presence of any articulation errors as measured in 1952 and in 1957 by the Wilson Articulation Test (*13*). After elimination of those groups in which only

two or three children were to be found, seven groups remained with sufficient numbers to warrant investigation:

Group I, children without articulation errors in 1952, no formal speech training (N = 8);

Group II, children with articulation errors in 1952, no speech lessons, no articulation errors in 1957 (N = 7);

Group III, children without articulation errors in 1952 following Wilson speech improvement lessons (N = 7);

Group IV, children with articulation errors in 1952 after Wilson speech improvement lessons, no additional speech therapy, no articulation errors in 1957 (N = 17);

Group V, children with articulation errors in 1952 after Wilson speech improvement lessons, no additional speech therapy, articulation errors still present in 1957 (N = 16);

Group VI, children with articulation errors in 1952 after Wilson speech improvement lessons, additional speech therapy (an average of 23 hours of public school group speech therapy), articulation errors still present in 1957 (N = 10);

Group VII, children with articulation errors in 1952, no formal speech training, articulation errors still present in 1957 (N = 29).

Factors. Measure of Maternal Personalities. The Gordon Personal Profile was selected to measure personality characteristics (five factors) of the mothers because of its brevity and test construction minimizing falsification. Previous research (7, pp. 12-15) established its acceptable validity and reliability. In the test four basic factors, Ascendency, Responsibility, Emotional Stability, and Sociability, are measured; the four are then combined to form a Total Gordon Scale on which low scores indicate poor adjustment. Local norms developed from the data gathered for this study (*1*) were used for comparisons between high and low scoring mothers.

Measure of Maternal Attitudes. Maternal attitudes were measured by the Parental Attitude Research Instrument (PARI) developed by Schaefer and Bell (*11*) and widely used for research purposes. The three weighted factors (*16*) utilized in this study included: Authoritarian-Control, Hostility-Rejection, and Democratic Attitudes.

Measure of Socioeconomic Level. A rating (one factor) of socioeconomic level on a six-point scale was made by the single interviewer utilizing the quality of the home, furnishings, and dwelling area as well as the occupation and education of the parents in the subjective judgment (*1*). Upper-class individuals in this study would correspond to Warner's (*6*, pp. 90-98) Upper and Upper-Middle classification, middle-class individuals to Warner's Lower-Middle and Upper-Lower classes, and lower-class individuals to Warner's Upper-Lower and Lower-Lower groupings.

Personal Data. Certain information concerning the family (five factors) was obtained through the use of a questionnaire. Items included were maternal education, paternal education, maternal age, maternal employment, and number of children.

Estimate of IQ. The IQ of the children, as obtained by Wilson (*13*), was the fifteenth factor used in the analysis.

Summary of 15 Factors. The 15 factors, thus, were as follows: maternal personality, five factors, measured by the Gordon Personal Profile; maternal attitude, three factors, measured by the Parental Attitude Research Instrument (PARI); family socioeconomic level, one factor, derived from interview, questionnaire, and occupational information from the schools; family background, five factors, based on infor-

mation from questionnaire; IQ, one factor, measured by Wilson study.

Interviewing Procedure. An interviewing procedure was used with the mothers. Their cooperation was obtained, the PARI and Gordon tests were explained, and judgments of socioeconomic status and maternal speech were made during the initial contact. Most test materials and the questionnaire were left with the mothers to be answered at their convenience. In three cases, individual administration of the materials was necessary due to reading difficulties. Materials were returned in stamped, self-addressed envelopes which were coded to assure identification. In order to reduce emotional reactions that are known (*2*) to bias maternal reports, the study was conducted under the guise of an opinion poll. There was no evidence that any mother was aware of a relationship between the study and the speech of her children.

Results

Correlations between the 15 factors studied indicated that there were high intercorrelations among the five Gordon scales and between the education and socioeconomic factors (*1*). Those correlations which are important to the interpretation of the study are given in Table 1. Analyses of variance, one for each of the 15 factors, were carried out to identify those factors on which groups differed significantly. Those with differences significant beyond the .01 level were Emotional Stability, Total Gordon, socioeconomic level, paternal education, and PARI Hostility-Rejection factors. Those with differences significant beyond the .05 level (but not at the .01 level) were Ascendency, Responsibility, Sociability,

TABLE 1. Correlations between the 10 significant factors: 1. PARI Hostility-Rejection (H-R); 2. Ascendency (Ascend); 3. Responsibility (Resp); 4. Emotional Stability (E Stab); 5. Sociability (Sociab); 6. Gordon Total Score (GTS); 7. Socioeconomic (Socioec); 8. Paternal Education (P Ed); 9. Maternal Education (M Ed); 10. Child's IQ (Ch IQ). (Means of the first six factors are weighted.)

Factor	*Means*	*SD*	*Correlations*								
			1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1. H-R	64.48	10.77									
2. Ascend	14.29	8.01	–.15								
3. Resp	22.38	9.11	–.23*	.54†							
4. E Stab	18.00	9.66	–.19*	.51†	.79†						
5. Sociab	16.71	8.78	–.22*	.88†	.55†	.44†					
6. GTS	41.51	23.73	–.28†	.84†	.76†	.80†	.81†				
7. Socioec	3.20	1.59	–.21*	.19*	.18*	.19*	.19*	.21*			
8. P Ed	12.18	5.02	–.15	.11	.12	.13	.10	.16	.54†		
9. M Ed	11.99	3.48	–.21*	.16	.18*	.15	.12	.19*	.32†	.75†	
10. Ch IQ	99.84	23.99	.02	.05	.10	.01	.04	.05	.32†	.33†	.23*

*Significant beyond the .05 level.
†Significant beyond the .01 level.

maternal education, and children's IQ. There were no significant differences between groups for maternal employment, maternal age, number of children, PARI Authoritarian-Control, and PARI Democratic Attitudes.

Differences between Groups. The 10 factors in which there were significant differences between groups were further analyzed by the Duncan (*5*) test utilizing the harmonic mean to correct for differences in group sizes. This analysis indicated which groups differed significantly (.05 level) from which other groups on each of the 10 significant factors. Results (*1*) are summarized below.

Group I scored significantly higher than at least four of the other groups in socioeconomic level, maternal education, Emotional Stability, and Total Gordon Scale. Likewise this group showed the least Hostility-Rejection, differing significantly from Groups III, VI, IV, and VII, and was significantly higher than three groups in paternal education.

Contrasting sharply with the Group I mothers of children having superior speech were the mothers of Group VI, whose children represented the most severe articulation problems in the study. Group VI was significantly lower than at least four other groups on Ascendency, Responsibility, Sociability, and Total Gordon Scale, and significantly lower than three other groups in socioeconomic level and Emotional Stability. Mothers in this group scored more hostile than three other groups as measured by the PARI Hostility-Rejection Scale. This group was significantly lower than at least one other group in paternal and maternal education and IQ of the child.

Although Group II was significantly lower than Group I in Emotional Stability, this was the only factor on which the two groups differed significantly from each other. A comparison of the success in articulation of these two criterion groups may suggest the negative effect of poorer maternal adjustment, and a comparison with Group III implies the positive value of kindergarten speech improvement lessons.

Group III revealed the greatest amount of Hostility-Rejection of all the groups, differing significantly from three other groups. Group III was significantly lower than Group I in maternal education, Responsibility, Emotional Stability, and the Gordon Total Scale. It was also significantly higher than Group VI on the Gordon Total Scale and significantly higher than two other groups in socioeconomic level and IQ of child. The speech success of the criterion group suggests that kindergarten speech improvement may be important as a means of overcoming possible bad effects of maternal Hostility-Rejection of home and family. It also suggests exploration of the possibility that superior intelligence might facilitate the early learning of errorless articulation.

Groups IV and V did not differ significantly from each other in any of the 10 factors although the members of criterion Group IV achieved errorless articulation at the time of this study while members of criterion Group V did not. Differences in sex distribution in the two criterion groups suggested this might be a relevant factor. Both groups were among the lower socio-

TABLE 2. Distribution of children according to participation in speech improvement lessons in 1952 (no lessons, lessons), socioeconomic level, and presence or absence of articulation errors in 1957 (error, no error).

Socioeconomic Level	*No Lessons*		*Lessons*	
	No Error	*Error*	*No Error*	*Error*
Upper	10	4	4	3
Middle	5	11	12	13
Lower	0	15	8	9

economic and educational levels of the seven groups studied and were significantly higher than Group VI in Emotional Stability, Total Gordon, Ascendency, and Responsibility.

Group VII was also among the lower socioeconomic and educational levels. This group was significantly lower than Group I in Emotional Stability and Total Gordon Scale, and significantly higher than Group VI in Ascendency and Responsibility.

It is to be noted that socioeconomic level, maternal adjustment, maternal attitudes rejecting home and family, and intelligence and sex of the children appeared to be of especial interest. In order to clarify the relative importance of these variables, further analysis of the over-all effect of each upon children with articulation problems compared with those without articulation errors was necessary.

Socioeconomic Level. Groups II and I were significantly higher in socioeconomic level than Groups VI, VII, IV, and V according to the Duncan test (*1*). Group III was significantly higher than Groups VI and VII.

In the upper socioeconomic levels 67% of the children were without articulation errors as compared with 41% in the middle classes and 25% in the lower socioeconomic levels. These differences were significant[2] beyond the .02 level ($\chi^2_{obs} = 8.49 > \chi^2_{.02} = 7.82$). Analysis of covariance of the PARI Hostility-Rejection factor with the socioeconomic factor partialled out was completed to determine the independence of the PARI factor. This analysis yielded an F of 2.43 which showed differences between the groups significant beyond the .05 level. Since the correlation between socioeconomic level and each of the five Gordon scales was equally low (Table 1), it can be assumed that these variables are also relatively independent of socioeconomic factors.

Speech Improvement Lessons. After five years, children of lower socioeconomic groups who had been in the Wilson kindergarten speech improvement program had attained a level of articulation achievement approximating that of middle and upper socioeconomic group members (see Table 2). Children who received speech improvement lessons did not differ significantly in socioeconomic level from children who did not participate in the program.

Differences in articulation related to socioeconomic level among children

[2]Unless otherwise indicated, all tests of significance utilized the chi square or Fisher exact tests.

who had received speech training were not significant ($\chi^2_{obs} = 0.20$). Highly significant differences in articulation related to socioeconomic level ($\chi^2_{obs} = 16.5 > \chi^2_{.001} = 13.82$) were found among those children who had not received kindergarten lessons. As Table 2 indicates, none of the lower class children who did not receive speech training had errorless articulation after the five year period. Although individuals in the training and no-training groups were not significantly different in articulation error at the upper and middle socioeconomic levels, they did differ significantly in the lower socioeconomic group (.002 level of significance). The value of Wilson kindergarten lessons in improving the articulation of children in lower socioeconomic groups seems to be clearly substantiated.

Maternal Education. Half of the mothers having at least a high school education had children without articulation problems while only 23% of the mothers with less than a high school education had children with errorless articulation. This difference was significant beyond the .05 level ($\chi^2_{obs} = 4.44 > \chi^2_{.05} = 3.84$).

Maternal Adjustment. None of the mothers in Group I scored below the median of the Total Gordon Scale, and none in Group VI scored above the median of that scale ($\chi^2_{obs} = 7.12 > \chi^2_{.01} = 6.64$). In this study maternal scores on the Total Scale (which ranges from −30 to 36) of 34 and above and on the Emotional Stability Scale (ranging from −15 to 18) of 14 and above coincided with absence of articulation error in 80% of the cases. Among the mothers who scored −20 and below on the Gordon Total and −11 and below on Emotional Stability, 80% had children with articulation problems in 1957.

People scoring in the bottom 5% of the Gordon Total Scale have been shown to be poorly adjusted (7). A test of the hypothesis that children with more severe articulation errors tend to have mothers who are maladjusted is provided by an examination of the scores of mothers in Group VI. The presence of two low scorers in Group VI was not significant ($\chi^2_{obs} = 2.69 > \chi^2_{.20} = 1.64$). No significant relationship was found between the articulation of children and maternal scores which comprised the middle 85% on the Gordon Total Scale (see Table 3).

TABLE 3. Placement (with number of cases in each rank) of the seven groups (I, II, . . ., VII) on the Total Scale of the Gordon Personal Profile.

Rank	*No Error 1957*				*Error 1957*		
	I	*II*	*III*	*IV*	*V*	*VI*	*VII*
Very High (7%)	2		1	2			1
High (24%)	5	1	1	2	5		6
Average (38%)	1	4	1	7	9	1	11
Low (24%)		1	3	4	2	5	4
Very Low (7%)				1		2	3
Totals	8	6	6	16	16	8	25

TABLE 4. Distribution of children according to participation in speech improvement lessons in 1952 (no lessons, lessons), maternal Hostility-Rejection (H-R) scores, and presence or absence of articulation errors in 1957 (error, no error).

H-R Scores	*No Lessons*		*Lessons*	
	No Error	*Error*	*No Error*	*Error*
High	0	10	12	6
Average	9	11	6	10
Low	6	8	6	10

Twelve mothers of children having articulation errors scored high or very high on the Total Scale, and 10 mothers of children without articulation problems scored low or very low on this scale. While this might be attributed to the insensitivity of the Gordon Personal Profile, there is a possibility that only severe maternal maladjustment could have a measurable effect upon children's articulation.

Maternal Hostility-Rejection. The percentage of mothers scoring high, average, or low on the Hostility-Rejection factor did not differ significantly in mothers of children in the training and no-training groups (see Table 4). Differences in articulation related to maternal Hostility-Rejection scores were not significant ($\chi^2_{obs} = 3.92 > \chi^2_{.20} = 3.22$) among children who had received speech improvement training. Significant differences in articulation related to maternal Hostility-Rejection scores were found among children who had not received kindergarten speech improvement training ($\chi^2_{obs} = 6.55 > \chi^2_{.05} = 5.99$).

None of the 10 children in the no-training group whose mothers scored high in Hostility-Rejection had errorless articulation. Of the 18 children in the training group whose mothers scored high in Hostility-Rejection, 12 had errorless articulation. This difference was significant at the .0006 level. Differences in articulation between training and no-training groups were not significant among children whose mothers scored average or low in Hostility-Rejection. The value of Wilson kindergarten improvement lessons in

TABLE 5. Placement (with number of cases in each rank) of the seven groups of mothers (I, II, . . . , VII) on the Sociability Scale of the Gordon Personal Profile.

Rank	*No Error 1957*				*Error 1957*		
	I	*II*	*III*	*IV*	*V*	*VI*	*VII*
Very High (7%)		3	3		2		2
High (24%)	2			2	4		6
Average (38%)	5	2	2	8	6	2	9
Low (24%)	1	1	1	6	2	6	5
Very Low (7%)					2		3
Totals	8	6	6	16	16	8	25

improving the articulation of children whose mothers scored high in Hostility-Rejection seems indicated by these findings.

Maternal Sociability Scores. Table 5 indicates that mothers of children having the best articulation, the mothers in Group I, had average Sociability ratings. Mothers of children without articulation problems, however, did not differ significantly in their Sociability scores from mothers of children having articulation errors.

Sex of the Children. Approximately equal proportions of boys and girls in this study had articulation problems. Although Groups IV and V were highly similar in many characteristics, there were more girls than boys among the errorless members of Group IV. This difference was not significant.

Discussion

Since most children achieve errorless articulation by the age of eight (*9*), it was assumed that all of the children studied could have been in Group II, whose criterion members achieved errorless articulation by fifth grade without receiving formal speech training. However, some children were superior, attaining errorless articulation at kindergarten age, while others had gross articulation errors after receiving many hours of speech therapy. In the sense that any child could have been in any group, the groups served as controls for each other, and differences between groups provided possible explanations for the children's performance.

This study was limited by the consideration of only selected environmental factors and by the use of testing instruments lacking in refinement. It is suggested therefore that differences found in spite of these limitations might be considered seriously as guideposts for more thorough research with larger groups. Since only a few of the children studied had severe articulation problems, conclusions are perhaps more applicable to a public school than to a clinical population.

Effect of Kindergarten Speech Improvement. The value of Wilson's speech improvement training over a period of time appears to be indicated by the results of this study. The number of children having articulation errors was markedly greater in the lower socioeconomic classes than in the middle or upper classes among children who had not received kindergarten improvement lessons. Members of lower socioeconomic classes who had received improvement lessons achieved as well as their peers in the middle and upper class groups. From this it might be concluded that it would be valuable to conduct similar speech improvement programs in schools serving lower socioeconomic areas.

The importance of the effect of marital problems upon articulation development was suggested by Wood's (*14*) list of disruptive home influences. The PARI Hostility-Rejection factor provides a measure of hostility toward children and husband and a rejection of the maternal role. Women high in this factor were found to have needs for achievement outside of the home, to lack motivation for care of others, and to show aggressiveness in interpersonal relationships (*15*). Although mothers in Group III showed high Hostility-Rejection scores, their children achieved

errorless articulation following speech improvement lessons. Group II children, who did not receive improvement lessons, failed to achieve errorless articulation until later although their mothers provided an atmosphere which was accepting of home and family.

Further substantiation of the utility of speech improvement training is provided through consideration of all the children of mothers having high Hostility-Rejection scores. Children who had received speech improvement training had significantly greater success in articulation than did those in the no-training group. Possibly the Hostility-Rejection scale provides a measure of maternal behavior as well as of maternal attitudes, or improvement lessons can perhaps counteract effects presumed to be of psychological origin. Further investigation of what the Hostility-Rejection scale measures and of the possibility that improvement lessons can aid in correcting articulation errors originating in psychological processes is required to form a valid conclusion.

Measurement of Maternal Adjustment. In agreement with Wood's (*14*) findings, mothers of children having the best articulation development scored highest in adjustment ratings while mothers of children with articulation problems which were resistant to therapy scored lowest on the adjustment scale. Wood's (*14*) implication that mothers of children with articulation errors have a tendency toward maladjustment was not substantiated by this study. However, relationships in both studies suggest the possibility that mothers who are so poorly adjusted as to be in need of psychotherapy may have a detrimental effect on their children which might be evidenced by functional articulation errors.

In considering the possibility of maternal maladjustment in a diagnostic evaluation, it is helpful to establish cutting scores on personality tests in conjunction with speech achievement of children. The practicality of the Gordon Personal Profile for speech assessment was crudely evaluated by administering it to 19 mothers of children who had articulation errors and who were receiving therapy at the Purdue Speech and Hearing Clinic. Predictions of progress were made for children of the four mothers scoring beyond the cut-off points established in the main study. Progress evaluations given by the speech therapists confirmed the prediction in three of four cases. In the fourth case, the therapist commented on lacking insight into the causes of the articulation problem. Perhaps the greatest utility for evaluations of maternal adjustment would be in such cases of doubtful etiology.

Intelligence. There were significant differences between some of the groups as measured by the Goodenough Draw-a-Man test of intelligence administered by Wilson (*13*). These differences may have been related to the speech improvement of children in Group III and the failure of other children to improve. Since an IQ score obtained at kindergarten age by means of a nonverbal intelligence test is of restricted validity, further research would be required to determine whether or not superior intelligence permits earlier and more accurate articulation than is possible with average intelligence.

PARI Factors. The absence of significant differences between groups on the

Authoritarian-Control and Democratic Attitudes factors may be due to the lack of sensitivity of the PARI as a measuring instrument or to the absence of relationships between these maternal attitudes and children's speech. Moll's (*8*) evaluation of two maternal attitude measures revealed limited utility for the field of speech. Refinements of the Hostility-Rejection factor may provide a tool which would be useful in uncovering maternal feelings toward home and family, a factor which appears to have a more direct relationship to children's speech than other attitudes measured by the PARI.

Summary

This study investigated long term effects of a kindergarten speech improvement program upon articulation in relation to socioeconomic level, certain maternal personality traits, and maternal family attitude factors. Seven mutual-control groups were established according to children's articulation and history of formal speech training. The Gordon Personal Profile, Parental Attitude Research Instrument, and a personal data questionnaire were given to 94 mothers who were evaluated on socioeconomic level.

Results indicated that children in lower socioeconomic groups who participated in kindergarten speech improvement achieved articulation success approximating that of upper class groups. Improvement training also appeared to counteract negative effects upon children's articulation associated with high maternal Hostility-Rejection scores. A relationship seemed to exist between children's articulation and maternal scores in the extremes of a personality test, and cut-off scores were given. Effects upon articulation of maternal rejection, maternal adjustment, and of superior intelligence in the child appeared to warrant additional research.

Acknowledgments

The author is indebted to Dr. Betty Ann Wilson, Associate Professor of Speech and Education, Purdue University, for assistance in planning the research and for the use of her original data, and to Dr. Charles D. Smock, Associate Professor of Psychology, Purdue University, for his criticisms of the manuscript.

References

1. Andersland, Phyllis Burgess, Relationships between attitudes and personality traits of mothers and the articulation development of their children. Master's thesis, Purdue Univ., 1959.
2. Bell, R. Q., Retrospective attitude studies of parent-child relations. *Child Develpm.*, 29, 1958, 323-338.
3. Bernreuter, R. G., *The Personality Inventory: Manual.* Stanford, Calif.: Stanford Univ. Press, 1935.
4. Drexler, H. G., The elements of articulation skill: prediction of later ability from a test battery administered in kindergarten. Master's thesis, Purdue Univ., 1958.
5. Duncan, D. B., Multiple range and multiple F tests. *Biometrics*, 11, 1955, 1-42.
6. Eells, K., Davis, A., Havighurst, R. J., Herrick, V. E., and Tyler, R. W., *Intelligence and Cultural Differences.* Chicago: Univ. Chicago Press, 1951.
7. Gordon, L. V., *Gordon Personal Profile: Manual.* New York: World Book, 1953.
8. Moll, K., An investigation of the attitudes of mothers of articulatory-

defective and speech-retarded children. M.A. thesis, Univ. Iowa, 1959.

9. Poole, Irene, Genetic development of articulation of consonant sounds in speech. *Elem. Eng. Rev.*, 11, 1934, 159-161.
10. Powers, Margaret Hall, Functional disorders of articulation—symptomatology and etiology. In L. E. Travis (Ed.), *Handbook of Speech Pathology*. New York: Appleton-Century-Crofts, 1957.
11. Schaefer, E. S., and Bell, R. Q., Development of a parental attitude research instrument. *Child Develpm.*, 29, 1958, 339-361.
12. Tiegs, E. W., Clark, W, W., and Thorpe, L. P., *Manual of Directions: California Test of Personality—adult series*. Los Angeles: Calif. Test Bur., 1942.
13. Wilson, Betty Ann, The development and evaluation of a speech improvement program for kindergarten children. *J. Speech Hearing Dis.*, 19, 1954, 4-13.
14. Wood, K. S., Parental maladjustment and functional articulatory defects in children. *J. Speech Dis.*, 11, 1946, 255-275.
15. Zuckerman, M., and Oltean, Mary, Some relationships between maternal attitude factors and authoritarianism, personality needs, psychopathology, and self-acceptance. *Child Develpm.*, 30, 1959, 27-36.
16. Zuckerman, M., Ribback, Beatrice Barrett, Monashkin, I., and Norton, J. A., Normative data and factor analysis on the Parental Attitude Research Instrument. *J. cons. Psychol.*, 22, 1958, 165-171.

Adaptation Effect and Spontaneous Recovery in Stuttering Expectancy

MARYANN PEINS

Two explanations have been specifically advanced to account for the adaptation effect in stuttering: an analogy to experimental extinction advanced by Wischner (*18*) and a theory of anxiety deconfirmation proposed by Johnson (*10*).

Wischner (*17, 18*) has hypothesized that stuttering behavior is a 'learned anxiety reaction system,' and that reduction of anxiety reinforces stuttering behavior. Wischner sees an analogy between phenomena in stuttering behavior and certain phenomena reported in conditioning and learning experiments; for example, he hypothesizes that the adaptation effect may be likened to experimental extinction. For experimental extinction to occur, there must be a withdrawal of reinforcement. As Wischner (*18*, p. 60) states, however, it is most difficult to identify the reinforcing mechanism or mechanisms operating in stuttering behavior.

Johnson (*10*) uses a theory of 'anxiety deconfirmation' to explain the adaptation effect. He suggests (*10*, pp. 259-260) that:

> To the degree that the stuttering and its consequences, as experienced, turn out to be less fearful than anticipated, the stutterer's anxiety is deconfirmed. With less intense anxiety, he is subsequently less apprehensive about stuttering on the words he has to say, less concerned about avoiding the stuttering he does anticipate, and so he does less to avoid it and does it with less tension. For this essential reason there is adaptation; that is, more and more talking leads to less and less stuttering because the stutterer gradually learns that in general his expectations are exaggerated and that speaking is nearly always not as bad as he thinks it will be.

An explanation offered in the literature for the increase in or recovery of stuttering following adaptation is an analogy to the spontaneous recovery noted in conditioning and learning experiments (*9, 11, 13, 18*). This explanation probably follows from the analogy drawn between adaptation in stuttering and the learning phenomenon of experimental extinction. Thus, it would seem that the validity of the explanation of the occurrence of recovery in stuttering is only as sound as the validity of the analogy between adaptation and experimental extinction.

An important aspect of the procedure used in studies on adaptation and recovery of stuttering is the requirement that stuttering subjects read or

Maryann Peins (Ph.D., Pennsylvania State University, 1958) is Associate Professor of Speech and Director, Speech and Hearing Center, Southern Connecticut State College. This article is based on a doctoral dissertation completed under the direction of Dr. James V. Frick and is an adaptation of a paper presented at the 1958 American Speech and Hearing Association Convention, New York.

speak spontaneously aloud. The explanations offered for the occurrence of these two phenomena are thus not only necessarily based on oral performance but also, and more importantly, seem crucially dependent on such oral performance. Both Wischner's analogy to experimental extinction and Johnson's theory of anxiety deconfirmation demand the elicitation of overt stuttering. Thus, it would seem that a study in which adaptation and recovery are demonstrated during a nonoral performance by stutterers would serve as an experimental test of the two explanations advanced for adaptation and recovery of stuttering.

The present study had as its point of departure the expectancy adaptation phenomenon which was first reported by Wischner (*16*, *18*) who found that when stutterers silently read a passage five successive times at one sitting and underlined words on which they would expect to stutter if they were to read the material aloud, there was a progressive decrease in the number of words underlined on each reading of the passage. Wischner's investigation indicated that it was possible to obtain an adaptation effect in a nonoral performance by stutterers under a condition similar to that employed in stuttering adaptation studies.

The present investigation sought to determine whether adaptation and recovery would be found to occur in the expectancy paradigm used by Wischner and, if they did, to conclude that the explanations explicitly offered for their occurrence in stuttering are invalid. In addition, the experimental design offered an opportunity to obtain information on other aspects relevant to the expectancy paradigm not reported in the literature.

Procedure

Subjects. Subjects were 32 adult stutterers, aged 15 to 33 years. They were randomly assigned to the experimental and control groups. The mean age of the experimental group (N=16) was 22.2 years, the median age, 21 years; the mean age of the control group (N=16) was 21.8 years, the median age, 20.5 years. The subjects had received speech therapy for periods ranging from zero to 10 years. None had previously participated nor was currently participating in stuttering research studies.

Materials. Test materials were two factual, 200-word reading passages (to be referred to as Passage A and Passage B) which were equated by the Flesch (*5*, p. 6) formula in terms of 'reading ease,' and by Brown's (*2*) criteria in terms of word weights.

The experiment was conducted in a small, quiet room with the subject seated at a table opposite the experimenter who was the only listener present. There were two conditions: a pretest and a test.

Pretest Condition. In the pretest, to evaluate differences between the groups on frequency of stuttering in oral reading, each subject read aloud Passage A five successive times at a single session. On duplicate copies of the passage the experimenter drew a line through each word the subject was judged to have stuttered (a marking procedure used also in the test condition).

Interval between Conditions. To remove or minimize any influence which participation in one session might have on results of a succeeding one, an interval of 24 hours was maintained between the pretest and the test and between the four days of the test. Jamison (*8*, pp. 245-248) found that a 4.5-hour period was sufficient for a relatively complete recovery of the stuttering response after adaptation. Leutenegger (*12*) reported that the amounts of recovery of the stuttering response following successive time delay intervals of 20 minutes, one hour, or 24 hours did not differ significantly. Jones (*11*), to whose study the present one is similar, found that statistically significant spontaneous recovery was demonstrated with a 24-hour interval between adaptation trials.

Test Condition. Reading Passage B was used for both groups in the test condition. In the oral portions, stuttered words were marked by the experimenter as in the pretest.

At the beginning of each session for the experimental group, five mimeographed copies of the reading passage, stapled together, were placed face down on the table before the subject. Instructions, essentially the same as used in Wischner's (*16*, *18*) study, were read by the experimenter at the beginning of each of the four successive daily sessions, directing the subject to read silently the five copies of the passage (zero time interval between readings) and to underline all words on which he would expect to stutter if he were to read the passage before a group. On the fourth day, following that day's fifth silent reading, there was a one-minute interval and the subject was asked to read the passage aloud five successive times.

At the beginning of each session for the control group, a mimeographed copy of the reading passage in a manilla folder was placed on the table before the subject. Instructions, read by the experimenter at the beginning of each of the four daily sessions, required the subject to read aloud the passage five successive times with zero time interval between readings. On the fourth day, following the fifth reading, there was a one-minute time interval and the subject was then asked to read the passage aloud five additional times.

TABLE 1. Results of *t* tests for evaluating differences in experimental group (N=16) mean differences in number of words expected to be stuttered in readings 1 and 5 (R_1, R_5) on each of four days and mean differences in number of words stuttered in readings 6 and 10 (R_6, R_{10}) on day 4 (df=15).

Day	*Mean 1*	*Mean 2*	*Diff*	*r*	*t*	*Sign Level*
R_1 *and* R_5						
1	17.81	19.50	−1.69	.90	.87	.4
2	23.25	23.31	− .06	.96	.03	.9
3	26.56	25.87	.69	.97	.37	.8
4	26.68	26.62	.06	.95	.03	.9
R_6 *and* R_{10}						
4	29.00	19.18	9.82	.84	5.65	.001

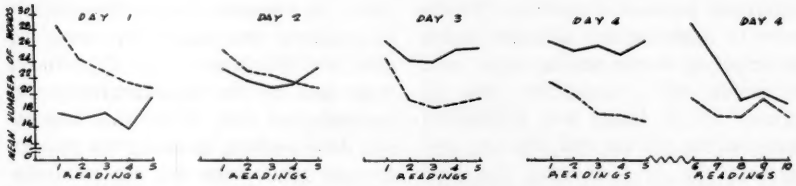

FIGURE 1. Mean frequencies of words expected to be stuttered and actually stuttered by the experimental group (solid line) and mean number of words stuttered by the control group (dashed line).

Observational Reliability of the Experimenter. Before the experiment, in order to determine her reliability in judging the occurrence of stuttering, the experimenter listened to a phonographically recorded sample of stuttering speech and at the same time followed a printed copy of the same material, marking the words judged to be stuttered. The same procedure was repeated a week later. By a formula employed by Tuthill (*15*), the experimenter's reliability was found to be .94, which compares favorably with Tuthill's values.

Results

Analysis of the frequency data (summarized in Table 1) by t tests[1] for correlated samples revealed no significant differences at the .05 level between the means of silent reading 1 and silent reading 5 on each of the four days for the experimental group. The difference between reading 6 and reading 10 on day 4 was significant at the .001 level, but these were oral readings. The differences between the means of reading 1 and reading 5 on each of four days for the control group were significant at the .05 level; however, the difference between reading 6 and reading 10 on day 4 was not significant.

The results of t tests ($df = 15$) indicated no significant differences at the .05 level between the mean number of words expected to be stuttered on each silent reading on each day (as well as the mean number of words stuttered on each oral reading on day 4) by the experimental group and the mean number of words stuttered on each reading by the control group.

Figure 1 presents graphically the frequency data on the mean number of words expected to be and actually stuttered by the experimental group and the mean number of words stuttered by the control group. The silent reading curves for the experimental group show no significant adaptation trend, whereas the curves for the control group, in general, tend to become progressively lower and significantly so. Curves for the experimental group resemble neither the curve reported in Wischner's (*18*) investigation nor the negatively accelerated curves reported in typical stuttering adaptation studies (*11*, *18*) or evidenced by the control group in this study. On each day the experimental group curve is

[1]There should be some caution in interpreting t-test results inasmuch as a number of such tests have been performed with the same data.

irregular and nonidentical in form with the curve for the preceding day. However, each curve indicates a decrease from reading 1 to reading 2 and tends to show an increase from either the third or fourth reading to the fifth reading. The means of the first and of the last silent readings are successively higher from day to day.

The curve for the oral readings (readings 6 through 10, day 4) for the experimental group is quite dissimilar to the curves depicting the silent readings, and resembles somewhat the typical adaptation curve as reported in the literature (*11*, *18*).

Spontaneous Recovery. Results of t tests for correlated samples between reading 5 of one day and reading 1 of the following day indicated that for the experimental group the difference for days 1 and 2 was significant at the .05 level ($t = 2.48$; $df = 15$), but the differences for days 2 and 3 and days 3 and 4 were not significant. For the control group, the only statistically significant increase was from reading 5, day 2, to reading 1, day 3. By inspection of Figure 1 the effect of the 24-hour interval between successive days for the experimental and control groups may be observed. In both groups the mean was higher on reading 1 of any day than on reading 5 of the preceding day; however, the increase became progressively less each day.

Effect of Previous Silent Inspections on Overt Stuttering. On the first oral reading trial (reading 6, day 4) for the experimental group the mean (Table 1) increased to 29.00, which was the largest mean for any reading, silent or oral, and approximates the mean of 28.50 for the control group on the first oral reading of day 1. A comparison of the performance of the experimental and control groups on any of the same oral readings indicated no significant difference at the .05 level between the mean frequencies of words stuttered by the two groups.

Discussion

The present study did not demonstrate an adaptation effect in the expectancy paradigm; rather, the experimental group evidenced an increase in mean frequency of stuttering expectancy (words expected to be stuttered) over the course of the four days of the experiment. Since an adaptation effect was not demonstrated, it is debatable whether the increase in frequency of expectancy on the first reading trial of each day should be regarded as spontaneous recovery of the expectancy phenomenon.

No conclusions can be drawn concerning the validity of the explanations explicitly advanced for adaptation and spontaneous recovery in stuttering since data were not yielded on the expectancy paradigm which might have invalidated these hypotheses.

Inasmuch as the present study does not provide evidence to substantiate Wischner's (*16*) findings, a question arises as to the reality or the replicability of the expectancy adaptation effect. This question can be resolved only by further research and, to this end, it is suggested that the following factors be considered in future studies of the expectancy adaptation effect: (a) age of subjects; (b) difference in procedure; and (c) degree of subjects' sophistication with respect to

previous participation in stuttering research studies.

Wischner's (*16*) subjects were younger stutterers (aged 12 to 26 years; mean age, 18, and median age, 16 years) than those used in the present experiment. Contradictory evidence in the literature (*4*, *6*) regarding age of subjects and adaptation does not resolve the question of whether age of subjects is a relevant variable in expectancy adaptation studies.

A difference in procedure may have contributed to the different results obtained in Wischner's (*16*) and the present experiment. Wischner conducted his experiment when all the subjects were together in a daily group meeting; in the present study, the subjects were seen individually by the experimenter at each session. Instructions (essentially the same as used in Wischner's study), read to the subjects at the beginning of each session, stated in part:

> . . . Think of the situation as that of standing and reading a passage before a group . . . Underline each word on which you think you will stutter should you be called upon to read the material orally before a group.

A daily increase in situational anxiety might be reflected in the results of the present study: the subjects, seen individually, might have been apprehensive about reading aloud before a group at the conclusion of any day's session. Some evidence for this is seen in the fact that the mean expectancy of stuttering decreased each day on readings 1, 2, and 3, and then increased on readings 4 and 5. In Wischner's experiment, there may have been less opportunity for situational anxiety: the subjects, being in a group situation, may have experienced a feeling of 'safety in numbers' and assumed it was highly improbable that each subject would be called upon to read before a group at the conclusion of the experiment.

There was a difference also, in the two studies, in the degree of subjects' sophistication with respect to previous participation in stuttering research. Wischner's subjects were clinically sophisticated; that is, within a relatively short period they served in many different adaptation studies which required them to read aloud passages and word lists. Wischner (*18*, pp. 27-28) suggests that there may have been 'interaction effects resulting from this situation.' In the present study, the subjects were clinically naïve; that is, none had previously participated nor were currently participating in stuttering research studies.

Reactive Inhibition and Nonpropositionality. Two other possible explanations for the adaptation effect are reactive inhibition and nonpropositionality of reading material. These were not involved in the design of the present study because they do not seem to depend upon an oral performance as do experimental extinction and anxiety deconfirmation.

Reactive inhibition, according to Hull (7, pp. 278-300), occurs with every response, whenever conditioned reactions, reinforced or nonreinforced, are evoked. Reactive inhibition is thought of as a negative drive state; as a given response is repeated, the response is not extinguished but only inhibited. Reference to reactive inhibition has been made by Wischner (*18*, p. 101) who proposes an experiment

on stuttering adaptation which would center around the amount of energy expended in stuttering and Luper (*14*, p. 9) who suggests that adaptation may be explained by reactive inhibition.

Reactive inhibition holds promise as a concept by which to account for adaptation in stuttering: it is a condition or state which builds up with the occurrence of each response, and, its rate of development is hastened by the massing of trials or response evocation. Inasmuch as the basic condition for the adaptation effect is the repeated reading of the same material with little, if any, time interval between successive readings, the hypothesized nature of reactive inhibition seems to provide a defensible explanation for the adaptation effect.

Reactive inhibition may be also a more parsimonious explanation for the adaptation effect than experimental extinction because the latter, as usually defined, is dependent upon the absence of reinforcement. Since reactive inhibition presumably accumulates with the occurrence of every response, whether reinforced or not, it would no longer be necessary to show that reinforcement was not present during the adaptation—a problem which has plagued Wischner and other investigators who have followed his analogy between adaptation and experimental extinction. Wischner (*18*, pp. 60-61) says: 'a most perplexing problem of stuttering behavior is that of identifying the reinforcing state of affairs.' Along similar lines, Jones (*11*, p. 50) states:

> The problem of just what is the reinforcement that is presumably being withheld in the case of stuttering adaptation has been of considerable interest and speculation.

A third example is from Leutenegger (*13*, p. 45) who says:

> In terms of Pavlovian experimental extinction it is difficult to determine the nature of the reinforcement, the diminishing of which would account for the adaptation phenomenon.

Since reactive inhibition is assumed to dissipate during the intervals between responding, the concept might also account for the phenomenon of recovery of stuttering after adaptation.

The other possible explanation for the adaptation effect in stuttering is that the material read aloud by the stutterer becomes nonpropositional or meaningless in nature with successive readings. Jones (*11*, pp. 1, 2) mentions that 'decrease, with successive readings of a given passage in the propositionality or communicative significance of the words stuttered' may possibly underlie the adaptation effect. 'It is a factor that seems to warrant more consideration than it has received. . . .' Quite possibly, with successive readings of a passage, the words may not hold 'anxiety value' or serve as 'noxious cues' for the stutterer, resulting in a progressive reduction in stuttering frequency.

Bloodstein (*1*) states that stutterers reported that reductions in stuttering occurred typically with utterance of material of low propositionality (nonsense material, isolated words, memorized recitations), and he further states that reduced communicative responsibility may have contributed to the reduction in stuttering frequency which stutterers reported in oral reading. Eisenson and Horowitz (*3*) found the frequency of stuttering to be

related to the meaningfulness of speech content.

It is hypothesized, therefore, that as the stutterer repeatedly reads the same passage, the increasingly nonpropositional nature of the material leads to a progressive decrease in frequency of stuttering. This hypothesis is not being advanced as the sole explanation for the adaptation effect; on the contrary, the adaptation effect might well be the result of the combined action of nonpropositionality and reactive inhibition.

Effect of Previous Silent Inspections on Overt Stuttering. Wischner (*18*) found that two rehearsals (inspections) of material, accompanied by deliberate or voluntary lip movements in forming the words, seemed to effect a reduction in frequency of overt stuttering on the first oral reading of such material. His findings prompted the question as to whether the same or similar results could be obtained with repeated silent inspections of material unaccompanied by deliberate or voluntary lip movements as the material was silently read.

The present study revealed that 20 silent inspections of material did not effect a decrease in frequency of overt stuttering on the first oral reading of such material. The fact that the subjects did not use 'deliberate or voluntary lip movements in forming words' may account for this finding. An adaptation effect was demonstrated to occur following 20 silent inspections of the same material; however, it cannot be concluded that this reduction in frequency of overt stuttering is attributable to the effect of repeated silent inspections.

Summary

The purpose of this study was to determine whether adaptation and recovery occur in the expectancy paradigm.

An experimental group of 16 stutterers silently read a passage five times on each of four consecutive days and marked words they would expect to stutter on oral reading; a control group of 16 stutterers orally read a passage five times on each of four consecutive days. Both groups then read the passage aloud five successive times.

An expectancy adaptation effect was not demonstrated, and thus no conclusions could be drawn concerning the validity of the explanations previously advanced for adaptation and recovery in stuttering. Evidence was obtained for spontaneous recovery within the expectancy paradigm only between days 1 and 2.

References

1. Bloodstein, O., Hypothetical conditions under which stuttering is reduced or absent. *J. Speech Hearing Dis.*, 15, 1950, 142-153.
2. Brown, S. F., The loci of stutterings in the speech sequence. *J. Speech Dis.*, 10, 1945, 181-192.
3. Eisenson, J., and Horowitz, Esther, The influence of propositionality on stuttering. *J. Speech Dis.*, 10, 1945, 193-197.
4. Falck, F. J., Interrelationships among certain behavioral characteristics, age, sex, and duration of therapy in a group of stutterers. Ph.D. dissertation, Pennsylvania State Univ., 1955.
5. Flesch, R., *How to Test Readability*. New York: Harper, 1951.
6. Harris, W. E., Studies in the psychology of stuttering. XVII, A study of the transfer of the adaptation effect in stuttering. *J. Speech Dis.*, 7, 1942, 209-221.
7. Hull, C. L., *Principles of Behavior, an Introduction to Behavior Theory*. New

York, London: D. Appleton-Century, 1943.

8. Jamison, Dorothy, J., Spontaneous recovery of the stuttering response as a function of the time following adaptation. In W. Johnson (Ed.), *Stuttering in Children and Adults*. Minneapolis: Univ. Minnesota Press, 1955.
9. Johnson, W., (Ed.), *Stuttering in Children and Adults*. Minneapolis: Univ. Minnesota Press, 1955.
10. Johnson, W., Brown, S. F., Curtis, J. F., Edney, C. W., and Keaster, Jacqueline, *Speech Handicapped School Children*. New York: Harper, 1956.
11. Jones, E. L., An investigation of stuttering viewed as learned behavior with special reference to experimental extinction and spontaneous recovery. Ph.D. dissertation, Univ. Iowa, 1950.
12. Leutenegger, R. R., Adaptation and recovery in the oral reading of stutterers. *J. Speech Hearing Dis.*, 22, 1957, 276-287.
13. Leutenegger, R. R., A study of adaptation and recovery in the oral reading of stutterers. Ph.D. dissertation, Univ. Iowa, 1953.
14. Luper, H. E., The consistency of selected aspects of behavior in the repetitions of stuttered words. Ph.D. dissertation, Ohio State Univ., 1954.
15. Tuthill, C. E., A quantitative study of extensional meaning with special reference to stuttering. *Speech Monogr.*, 13, 1946, 81-98.
16. Wischner, G. J., An experimental approach to expectancy and anxiety in stuttering behavior. *J. Speech Hearing Dis.*, 17, 1952, 139-154.
17. Wischner, G. J., Stuttering behavior and learning: a preliminary theoretical formulation. *J. Speech Hearing Dis.*, 15, 1950, 324-335.
18. Wischner, G. J., Stuttering behavior and learning: a program of research. Ph.D. dissertation, Univ. Iowa, 1947.

Call for Program Suggestions and Papers, Thirty-seventh Annual Convention of the American Speech and Hearing Association

NOVEMBER 5, 6, 7, and 8

HOTEL SHERMAN, CHICAGO, ILLINOIS

The ASHA Convention Program Committee announces a Call for Program Suggestions and Papers for the 1961 National Convention. Contributors are urged to submit abstracts and other program proposals immediately to insure that they may be given consideration during the planning stages for the Convention.

The Program Committee will continue the efforts of past Program Committees to upgrade the caliber of the Convention presentations. Program proposals will be carefully screened. The time available for the program will be less and the number of presentations to be scheduled will be fewer for the 1961 Convention than for the 1960 Convention. Contributors must supply sufficient information to the Program Committee to enable it to make judicious choices from among the proposals and to permit intelligent scheduling of the presentations selected. Since the program abstracts and summaries will be published prior to the Convention, they must be complete and in suitable publication form.

The organization of a program for a national convention is necessarily complex. Scrupulous adherence to both deadlines and forms for submitting program proposals is necessary if the Committee is to work effectively.

The Program Committee urges your active participation in planning the program. Any suggestions will be given careful consideration. If you wish to make a program proposal of any type, please read carefully the instructions given below.

Types of Programs and Sessions

In the interest of serving the many and varied needs of the membership, a number of different types of meetings and sessions will be scheduled. For the purposes of clarification and uniformity of listings, the following descriptions and definitions are provided.

Papers. Papers may deal with experiments, investigations, theoretical formulations, clinical practices, philosophic considerations, and policy or procedures in the various settings of our professional practice. Papers from related disciplines will be considered.

Symposium. A symposium is a meeting or series of meetings involving the presentation of two or more formal papers on a given topic.

Panel. A panel presentation is an informal discussion of a given topic by a group of five or six experts. There are no speeches per se and the Chairman interjects just enough ideas to set the stage and to give direction to the discussion.

Forum. There are two types of forum: an open forum, which involves extended audience discussion of a topic that has been briefly outlined; and a panel-forum, in which a panel is used for defining the topic and participating in the discussion. In either type, a question period, with special interrogators or selected written questions, may be used to guide discussion toward important implications.

Workshop. A workshop provides a structured learning situation in which the audience participates with the aid of resource persons.

Qualifications for Participation

Both members and nonmembers may submit proposals and participate in the program.

In order to insure maximum participation by the membership, the Executive Council has ruled that each individual be limited to two appearances on the program. This ruling has been interpreted not to apply to those instances in which individuals are invited to participate or where they are chairmen of sections.

Requirements for Proposals

All proposals for the program, inincluding abstracts, must be submitted by April 1, must be typed in quadruplicate, and may not exceed 300 words in length. All communications concerning the program are to be addressed as follows:

D. C. Spriestersbach, Program Committee Chairman
Department of Speech Pathology and Audiology
University of Iowa
Iowa City, Iowa

Obligations of Contributors

Persons accepted to appear on the program thereby assume the responsibility to appear in person as scheduled and to adhere rigidly to the time limits assigned. Participants must agree that no substantial changes will be made in accepted proposals without advance approval of the Program Committee. Should unforeseen contingencies arise which preclude the appearance as planned, the contributor must supply a substitute by advance arrangement with the section chairman.

Contributors assume the further responsibility to cooperate with the abstracts editor by meeting such deadlines as he shall indicate in making any revisions of abstracts and summaries deemed necessary. The Program Committee wishes to insure uniform coverage and quality of the published abstracts and summaries.

Form for Submitting Proposals

Identifying Data. These data are required of all proposals:

1. Title of paper or program.
2. Name and advanced degree of each participant.
3. Name of the institution or laboratory with which each participant is affiliated.
4. Designation of work which was done to satisfy the requirements for an advanced degree and identification of the director of the study.

5. Identification of work which was done as part of a sponsored research project with name of the sponsoring agency and the project code number.

6. Time required for presentation.

7. Audio-visual equipment needed or special physical setting required. Participants are requested to consider carefully the advantages and disadvantages of duplicated handouts of graphic and tabular material over presentations by slide projectors. However, projectors will be furnished if requested.

Form for Specific Types of Proposals. Contributed Papers. Research papers are proposed by submitting an abstract which contains a concise statement of the problem, procedures, and results. Results and conclusions should be stressed. All data should have been collected and the analysis completed at the time that the abstract is submitted. Papers concerned with data which have not been completely analyzed by April 1 will be considered only by special arrangement with the Program Chairman.

Papers dealing with matters other than research are proposed by submitting a summary of the subject matter to be covered. The summary should include a statement of the hypotheses being examined and the conclusions and implications to be drawn from the material.

Invited Papers. Members are urged to suggest names of eminent scholars, researchers, or professional workers who could make a valuable contribution to the program. However, invitations to appear are extended only by the Program Committee.

Symposia, Panels, Forums, and Workshops. Members may suggest proposals for symposia, panels, forums, and workshops. All participants should be identified as indicated in *Identifying Data* section above. Proposals should include a statement of the purpose and significance of the presentation, and a description of the subject matter to be covered. The specific contribution of each participant should be indicated.

Film Theater. Suggestions concerning films appropriate for the Film Theater should be sent directly to the Chairman of the Film Theater Subcommittee, Dr. Robert J. Duffy, 2511 Hazelwood Way, Palo Alto, California.

Scientific Exhibits. Free space for scientific exhibits will be available. Applications for Space Forms may be obtained by writing the Executive Secretary, American Speech and Hearing Association, 1001 Connecticut Avenue, N. W., Washington 6, D. C.

The Program Committee: Stanley H. Ainsworth, Walter W. Amster, Robert C. Bilger, H. Harlan Bloomer, John L. Boland, William E. Castle, David R. Dickson, Robert J. Duffy, Sue Earnest, James P. Egan, Aubrey Epstein, D. Robert Frisina, Robert Goldstein, Elise S. Hahn, James C. Hardy, Earl R. Harford, Helen S. Knight, George H. Kurtzrock, Herold S. Lillywhite, James E. McLean, Otto J. Menzel, June Miller, Sheila G. Morrison, Gene R. Powers, William F. Prather, Maryjane Rees, Josephine Simonson, Courtney Stromsta, Wayne L. Thurman, John C. Webster, Richard A. Winchester, Jozef J. Zwislocki, and D. C. Spriestersbach, Chairman.